

92394

STIC-Biotech/ChemLib

From: Spector, Lorraine
Sent: Friday, April 25, 2003 8:06 AM
To: STIC-Biotech/ChemLib
Subject: SEARCH request for Serial No. 09/674377

STIC,

Please search SEQ ID NO:1

- pending
- issued
- commercial

Thanks.

Lorraine Spector
703-308-1793
U.S. Patent and Trademark Office
Art Unit 1647
lorraine.spector@uspto.gov
CM1-10B11
Mailbox 10-B19

Edward Hart
Technical Info. Specialist
STIC/Biotech
CMI 6B02 Tel: 305-9203

Searcher: _____
Phone: _____
Location: _____
Date Picked Up: 4/25/03
Date Completed: 4/26/03
Searcher Prep/Review: _____
Clerical: _____
Online time: _____

TYPE OF SEARCH:
NA Sequences: _____
AA Sequences: 1
Structures: _____
Bibliographic: _____
Litigation: _____
Full text: _____
Patent Family: _____
Other: _____

VENDOR/COST (where applic.)
STN: _____
DIALOG: _____
Questel/Orbit: _____
DRLink: _____
Lexis/Nexis: _____
Sequence Sys.: 72
WWW/Internet: _____
Other (specify): _____

THIS PAGE BLANK (USPTO)

Spector, Lorraine

From: Spector, Lorraine
Sent: Friday, April 25, 2003 8:06 AM
To: STIC-Biotech/ChemLib
Subject: SEARCH request for Serial No. 09/674377

STIC,

Please search SEQ ID NO:1

-pending

-issued

-commercial

Thanks.

Lorraine Spector
703-308-1793
U.S. Patent and Trademark Office
Art Unit 1647
lorraine.spector@uspto.gov
CM1-10B11
Mailbox 10-B19

THIS PAGE BLANK (USPTO)

BioTech-Chem Library

Search Results

Feedback Form (Optional)



Scientific & Technical Information

The search results generated for your recent request are attached. If you have any questions or comments (compliments or complaints) about the scope or the results of the search, please contact *the BioTech searcher* who conducted the search *or contact*:

Mary Hale, Supervisor, 30
CM-1 Room 1E01

Voluntary Results Feedback Form

➤ *I am an examiner in Workgroup: 1647 (Example: 1610)*

➤ *Relevant prior art found, search results used as follows:*

- ☐ 102 rejection
- ☐ 103 rejection
- ☐ Cited as being of interest.
- ☐ Helped examiner better understand the invention.
- ☐ Helped examiner better understand the state of the art in their technology.

Types of relevant prior art found:

- ☐ Foreign Patent(s)
- ☐ Non-Patent Literature
(journal articles, conference proceedings, new product announcements etc.)

➤ *Relevant prior art not found:*

- ☐ Results verified the lack of relevant prior art (helped determine patentability).
- ☐ Search results were not useful in determining patentability or understanding the invention.

Other Comments:

Drop off completed forms at the Circulation Desk CM-1, or send to Mary Hale, CM1-1E01 or mary.hale@uspto.gov

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 25, 2003, 15:44:58 ; Search time 77 Seconds
(without alignments)
773.546 Million cell updates/sec

Title: US-09-674-377B-1
Perfect score: 2604
Sequence: 1 ERKRRNTIHEPKSAKTLI.....IPWDYCPISRCGDTTPTIV 447

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_101002.*

- 1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
- 2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
- 3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
- 4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
- 5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
- 6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
- 7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
- 8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
- 9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
- 10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
- 11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
- 12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
- 13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
- 14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
- 15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
- 16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
- 17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
- 18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
- 19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
- 20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
- 21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
- 22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
- 23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2601	99.9	697	21	AA198485 Hepatocyte growth
2	2601	99.9	697	21	AA198485 Sequence of a pept
3	2601	99.9	697	22	AAU04275 Nuclear ligand Pep
4	2601	99.9	697	22	AAU04275 Nuclear acid trans
5	2601	99.9	728	13	AA200005 Human hepatocyte g
6	2601	99.9	728	13	AA200005 Recombinant human
7	2601	99.9	728	14	AA200005 Competitive inhibi
8	2601	99.9	728	14	AA200005 Competitive inhibi
9	2601	99.9	728	14	AA200005 Vascular endotheli
10	2601	99.9	728	17	AAW00338 Human hepatic pare

11	2601	99.9	728	17	AAW00340 Wild type hepatocy
12	2601	99.9	728	17	AAW00340 Mutant hepatocyte
13	2601	99.9	728	17	AAW00340 Mutant hepatocyte
14	2601	99.9	728	17	AAW00340 Mutant hepatocyte
15	2601	99.9	728	17	AAW00340 Mutant hepatocyte
16	2601	99.9	728	17	AAW00340 Mutant hepatocyte
17	2601	99.9	728	19	AAW59922 Human leukocyte-de
18	2601	99.9	728	19	AAW59922 Human hepatocyte g
19	2601	99.9	728	19	AAW59922 Human hepatocyte g
20	2601	99.9	728	19	AAW59922 Human hepatocyte g
21	2601	99.9	728	21	AAW59922 Human hepatocyte g
22	2601	99.9	728	21	AAW59922 Human hepatocyte g
23	2601	99.9	728	21	AAW59922 Human hepatocyte g
24	2601	99.9	728	21	AAW59922 Human hepatocyte g
25	2601	99.9	728	21	AAW59922 Human hepatocyte g
26	2601	99.9	728	21	AAW59922 Human hepatocyte g
27	2601	99.9	728	21	AAW59922 Human hepatocyte g
28	2601	99.9	728	21	AAW59922 Human hepatocyte g
29	2601	99.9	728	21	AAW59922 Human hepatocyte g
30	2601	99.9	728	21	AAW59922 Human hepatocyte g
31	2601	99.9	728	21	AAW59922 Human hepatocyte g
32	2601	99.9	728	21	AAW59922 Human hepatocyte g
33	2601	99.9	728	21	AAW59922 Human hepatocyte g
34	2601	99.9	728	21	AAW59922 Human hepatocyte g
35	2601	99.9	728	21	AAW59922 Human hepatocyte g
36	2601	99.9	728	21	AAW59922 Human hepatocyte g
37	2601	99.9	728	21	AAW59922 Human hepatocyte g
38	2601	99.9	728	21	AAW59922 Human hepatocyte g
39	2601	99.9	728	21	AAW59922 Human hepatocyte g
40	2601	99.9	728	21	AAW59922 Human hepatocyte g
41	2601	99.9	728	21	AAW59922 Human hepatocyte g
42	2601	99.9	728	21	AAW59922 Human hepatocyte g
43	2601	99.9	728	21	AAW59922 Human hepatocyte g
44	2601	99.9	728	21	AAW59922 Human hepatocyte g
45	2601	99.9	728	21	AAW59922 Human hepatocyte g

ALIGNMENTS

RESULT 1	
AA198485	
ID	AA198485 standard; Protein; 697 AA.
XX	AA198485
AC	AA198485
DT	31-JUL-2000 (first entry)
DE	Hepatocyte growth factor Pep 20 used in nucleic acid transporter system.
XX	
KW	Transporter system; nucleic acid delivery; gene therapy; cancer;
KW	carcinogenesis; cardiovascular disease; infection.
XX	
OS	Synthetic.
XX	
PN	US6033884-A.
XX	
PD	07-MAR-2000.
XX	
PF	14-DEC-1993; 93US-0167641.
XX	
PR	20-MAR-1992; 92US-0855389.
PR	19-MAR-1993; 93WO-US02725.
XX	
XX	(BAYU) BAYLOR COLLEGE MEDICINE.
PI	Gottchalk S, Sparrow J, Cristiano RJ, Woo SLC, Smith LC;
XX	
DR	WPI; 2000-281993/24.
XX	
PT	System for transporting nucleic acid into cells, useful e.g. in gene
PT	therapy and for generating transgenic animals, comprises binding agent
PT	linked to nucleic acid, surface ligand and lytic agent

XX Disclosure; Figure 23A; 108pp; English.

XX The present invention relates to a transporter system for delivering

XX nucleic acid to a cell. The system comprises a nucleic acid binding

XX complex, consisting of a binding molecule bonded non-covalently to the

XX nucleic acid, and covalently to a surface ligand, and a lytic agent. The

XX binding molecule is spermine or a spermidine derivative. Nucleotide

XX sequences AAA36633-A36652 and peptide sequences AAY98456-Y98500 are used

XX in the construction of the transporter system of the invention. The

XX transporter system is used in gene therapy, particularly to deliver

XX nucleic acids to hepatocytes, muscle cells or bone forming cells, e.g

XX for treating cardiovascular disease, cancer, and infection. The

XX transporter systems are also used to create transgenic animals (as models

XX for human carcinogenesis or disease or for drug testing). Other uses

XX include transforming cells to produce proteins, or transfecting cells in

XX vitro to study the function of the nucleic acid. The use of a surface

XX ligand allows specific targeting of selected cells and tissues. The lytic

XX agent provides for release of the nucleic acid into the cellular

XX interior, from endosomes, without requiring endosomal or lysosomal

XX degradation.

XX Sequence 697 AA;

XX

PS Query Match 99.9%; Score 2601; DB 21; Length 697;

PS Best Local Similarity 99.8%; Pred. No. 1.3e-165;

PS Matches 446; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ERKRRNTIHEFKSAKTLIKIDPALKIKTKKVTADQCANRCKTRNKGLPFTCKAFVFDK 60

DB 1 QKRRNTIHEFKSAKTLIKIDPALKIKTKKVTADQCANRCKTRNKGLPFTCKAFVFDK 60

QY 61 ARKQCLWFFPNSMSGVKKKEFGHEFDLYENKDYIRNCIIIGKGRSYKGTVSIKSGIKCOP 120

DB 61 ARKQCLWFFPNSMSGVKKKEFGHEFDLYENKDYIRNCIIIGKGRSYKGTVSIKSGIKCOP 120

QY 121 WSSMIPHEHSFLPSSYRGKDLQENYCRNPRGEGGPGWCFTSNPEVRYEVCIDIPOCSEVEC 180

DB 121 WSSMIPHEHSFLPSSYRGKDLQENYCRNPRGEGGPGWCFTSNPEVRYEVCIDIPOCSEVEC 180

QY 181 MTCNGESYRGKDLQENYCRNPRGEGGPGWCFTSNPEVRYEVCIDIPOCSEVEC 240

DB 181 MTCNGESYRGKDLQENYCRNPRGEGGPGWCFTSNPEVRYEVCIDIPOCSEVEC 240

QY 301 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 360

DB 301 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 360

QY 361 YRGKNGKYNGLNLSQTRSLGTCSDMKKEDLHRHIFWEPDASKLENYCRNPDGDAHGPW 420

DB 361 YRGKNGKYNGLNLSQTRSLGTCSDMKKEDLHRHIFWEPDASKLENYCRNPDGDAHGPW 420

QY 421 CYTGNPLIPWDYCPISRCGEGTPTIV 447

DB 421 CYTGNPLIPWDYCPISRCGEGTPTIV 447

RESULT 2

AAAY59030

ID AAY59030 standard; peptide: 697 AA.

XX

AC AAY59030;

XX

XX 07-MAR-2000 (first entry)

DE

XX Sequence of a peptide ligand Pep20.

XX

XX Nucleic acid transport system; NTS; cell surface receptor; cytolysis;

XX nuclear membrane; lysis moiety; transgenic animal; human disease;

KW

XX nucleic acid delivery; cancer.

OS Synthetic.

XX US5994109-A.

XX

PD 30-NOV-1999.

XX

XX 03-JUN-1995; 95US-0460890.

XX

XX 14-DEC-1993; 93US-0167641.

XX

XX 20-MAR-1992; 92US-0855389.

PR

XX 19-MAR-1993; 93WO-US02725.

PR

XX 14-DEC-1993; 93US-0167641.

XX

XX (BAYU) BAYLOR COLLEGE MEDICINE.

PA

XX Woo SLC, Cristiano RJ, Gottchalk S, Sparrow J, Smith LC;

PI

XX WPI; 2000-038262/03.

DR

XX Nucleic acid transport system, useful for creating transgenic animals

XX for assessing human disease such as cancer in an animal model

XX

PS Disclosure; Fig 23A; 107pp; English.

XX

XX The invention relates to a nucleic acid transport system (NTS) for

XX delivering nucleic acid into a cell. The NTS contains but is not limited

XX to 5 components: (a) the nucleic acid or a macromolecule to be delivered;

XX (b) a moiety that recognizes and binds to a cell surface receptor or

XX antigen or is capable of entering a cell through cytolysis; (c) a nucleic

XX acid or macromolecular molecule binding moiety; (d) a moiety that is

XX capable of moving or initiating movement through a nuclear membrane; and/

XX or (e) a lysis moiety that enables the transport of the entire complex

XX from the cell surface directly into the cytoplasm of the cell. The NTS

XX delivers nucleic acid into the cellular interior as well as the nucleus

XX of specific cells. The NTS can be used to treat disorders by targeting

XX specific nucleic acid accordingly. The NTS can also be used to create

XX transgenic animals for assessing human disease, such as cancer, in an

XX animal model. The NTS can be used in vitro with tissue culture cells

XX which allows the role of various nucleic acids to be studied by targeting

XX specific expression into specifically targeted tissue culture cells. The

XX lysis agent within the NTS avoids the problem of endosomal/lysosomal

XX degradation.

XX

XX Sequence 697 AA;

XX

Query Match 99.9%; Score 2601; DB 21; Length 697;

Best Local Similarity 99.8%; Pred. No. 1.3e-165;

Matches 446; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ERKRRNTIHEFKSAKTLIKIDPALKIKTKKVTADQCANRCKTRNKGLPFTCKAFVFDK 60

DB 1 QKRRNTIHEFKSAKTLIKIDPALKIKTKKVTADQCANRCKTRNKGLPFTCKAFVFDK 60

QY 61 ARKQCLWFFPNSMSGVKKKEFGHEFDLYENKDYIRNCIIIGKGRSYKGTVSIKSGIKCOP 120

DB 61 ARKQCLWFFPNSMSGVKKKEFGHEFDLYENKDYIRNCIIIGKGRSYKGTVSIKSGIKCOP 120

QY 121 WSSMIPHEHSFLPSSYRGKDLQENYCRNPRGEGGPGWCFTSNPEVRYEVCIDIPOCSEVEC 180

DB 121 WSSMIPHEHSFLPSSYRGKDLQENYCRNPRGEGGPGWCFTSNPEVRYEVCIDIPOCSEVEC 180

QY 181 MTCNGESYRGKDLQENYCRNPRGEGGPGWCFTSNPEVRYEVCIDIPOCSEVEC 240

DB 181 MTCNGESYRGKDLQENYCRNPRGEGGPGWCFTSNPEVRYEVCIDIPOCSEVEC 240

QY 241 YTLDPHTRWEYCAIKTCADNTMNDTVPLETTCIOGEGYRGTVNTIWNIGIPCORWDS 300

DB 241 YTLDPHTRWEYCAIKTCADNTMNDTVPLETTCIOGEGYRGTVNTIWNIGIPCORWDS 300

QY 301 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 360

DB 301 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 360

QY 361 YRGKNGKYNGLNLSQTRSLGTCSDMKKEDLHRHIFWEPDASKLENYCRNPDGDAHGPW 420

DB 361 YRGKNGKYNGLNLSQTRSLGTCSDMKKEDLHRHIFWEPDASKLENYCRNPDGDAHGPW 420

QY 421 CYTGNPLIPWDYCPISRCGEGTPTIV 447

DB 421 CYTGNPLIPWDYCPISRCGEGTPTIV 447

RESULT 2

AAAY59030

ID AAY59030 standard; peptide: 697 AA.

XX

AC AAY59030;

XX

XX 07-MAR-2000 (first entry)

DE

XX Sequence of a peptide ligand Pep20.

XX

XX Nucleic acid transport system; NTS; cell surface receptor; cytolysis;

XX nuclear membrane; lysis moiety; transgenic animal; human disease;

KW

CC such as hormones, growth factors, clotting factors,
 CC apolipoproteins, receptors, drugs, oncogenes, tumor antigens, tumor
 CC suppressors, viral antigens, parasitic antigens, and bacterial antigens.
 CC The transporter system uses lysis agents to overcome the problems of
 CC endosomal/lysosomal degradation seen with prior art systems.

XX
 CC Sequence 697 AA;
 Query Match 99.9%; Score 2601; DB 22; Length 697;
 Best Local Similarity 99.8%; Pred. No. 1.3e-165;
 Matches 446; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ERKRRNTIHEFKSAKTLIKIDPALKIKTKVNTADOCANRCTRNKGLPFTCKAFVDEK 60
 :|||||
 Db 1 QKRRNTIHEFKSAKTLIKIDPALKIKTKVNTADOCANRCTRNKGLPFTCKAFVDEK 60
 :|||||
 QY 61 ARKQCLWFFPNSMSGVKKFEGHEFDLYENKDYIRNCIIIGKRSYKGTVSITKSGIKQCP 120
 :|||||
 Db 61 ARKQCLWFFPNSMSGVKKFEGHEFDLYENKDYIRNCIIIGKRSYKGTVSITKSGIKQCP 120
 :|||||
 QY 121 WSSMIPHEHSLPSSYRGKDLQENYCRNPRGEGGPGWCFSTSNPEVRYEVCIDIPOCSEVEC 180
 :|||||
 Db 121 WSSMIPHEHSLPSSYRGKDLQENYCRNPRGEGGPGWCFSTSNPEVRYEVCIDIPOCSEVEC 180
 :|||||
 QY 181 MTCNGESYRGKDLQENYCRNPRGEGGPGWCFSTSNPEVRYEVCIDIPOCSEVEC 240
 :|||||
 Db 181 MTCNGESYRGKDLQENYCRNPRGEGGPGWCFSTSNPEVRYEVCIDIPOCSEVEC 240
 :|||||
 QY 241 YTLDPHTRWEYCAIKTCADNTMNDTVPLETTETCIGQGEYRGVTNTWNGIPQORWDS 300
 :|||||
 Db 241 YTLDPHTRWEYCAIKTCADNTMNDTVPLETTETCIGQGEYRGVTNTWNGIPQORWDS 300
 :|||||
 QY 301 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFSTTDPNIRVGYCSQIPNCDMSHGQDC 360
 :|||||
 Db 301 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFSTTDPNIRVGYCSQIPNCDMSHGQDC 360
 :|||||
 QY 361 YRGNGKNYMGNLSQTRSGLTCSMWKDNMEDLHRHIFWEPDASKLNEYCRNPDDDAHGWP 420
 :|||||
 Db 361 YRGNGKNYMGNLSQTRSGLTCSMWKDNMEDLHRHIFWEPDASKLNEYCRNPDDDAHGWP 420
 :|||||
 QY 421 CYTGNPLIPWDYCPISRCGDTTPTIV 447
 :|||||
 Db 421 CYTGNPLIPWDYCPISRCGDTTPTIV 447
 :|||||

RESULT 5
 AAR20005
 ID AAR20005 standard; Protein; 728 AA.

XX AAR20005;
 AC
 DT 24-MAR-1992 (first entry)
 XX Human hepatocyte growth factor.
 DE HGF; kidney regeneration; nephritis.

XX Homo sapiens.

XX EP462549-A.

XX PD 27-DEC-1991.

XX PF 18-JUN-1991; 91EP-0109923.

XX PR 19-JUN-1990; 90JP-0158841.

XX PA (TOYM) TOYO BOSEKI KK.

XX PI Nakamura T;

XX DR WPI; 1992-000939/01.

XX N-PSDB; AAQ20049.

XX Agent contg. hepatocyte growth factor and carrier - used for
 PT treating renal diseases and promoting nephrocyte growth and as
 PT diagnostic for renal diseases
 XX
 PS Disclosure; Fig 1; 15pp; English.

XX Human HGF comprises an alpha-chain of 440 amino acids and a
 CC beta-chain of 234 amino acids. There are 4 kringle domains in the
 CC alpha-chain, similar to that of plasmin; the beta-chain has about
 CC 37 per cent homology with the beta-chain of plasmin having serine
 CC protease activity. Homology of the amino acid sequence of rat HGF
 CC and human HGF is 91.6 per cent in the alpha-chain and 88.9 per cent
 CC in the beta-chain. HGF has been found to be an agent for nephrocyte
 CC growth and is useful as a treatment for renal failure.

XX Sequence 728 AA;

Query Match 99.9%; Score 2601; DB 13; Length 728;
 Best Local Similarity 99.8%; Pred. No. 1.4e-165;
 Matches 446; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ERKRRNTIHEFKSAKTLIKIDPALKIKTKVNTADOCANRCTRNKGLPFTCKAFVDEK 60
 :|||||
 Db 32 QKRRNTIHEFKSAKTLIKIDPALKIKTKVNTADOCANRCTRNKGLPFTCKAFVDEK 91
 :|||||
 QY 61 ARKQCLWFFPNSMSGVKKFEGHEFDLYENKDYIRNCIIIGKRSYKGTVSITKSGIKQCP 120
 :|||||
 Db 92 ARKQCLWFFPNSMSGVKKFEGHEFDLYENKDYIRNCIIIGKRSYKGTVSITKSGIKQCP 151
 :|||||
 QY 121 WSSMIPHEHSLPSSYRGKDLQENYCRNPRGEGGPGWCFSTSNPEVRYEVCIDIPOCSEVEC 180
 :|||||
 Db 152 WSSMIPHEHSLPSSYRGKDLQENYCRNPRGEGGPGWCFSTSNPEVRYEVCIDIPOCSEVEC 211
 :|||||
 QY 181 MTCNGESYRGKDLQENYCRNPRGEGGPGWCFSTSNPEVRYEVCIDIPOCSEVEC 240
 :|||||
 Db 212 MTCNGESYRGKDLQENYCRNPRGEGGPGWCFSTSNPEVRYEVCIDIPOCSEVEC 271
 :|||||
 QY 241 YTLDPHTRWEYCAIKTCADNTMNDTVPLETTETCIGQGEYRGVTNTWNGIPQORWDS 300
 :|||||
 Db 272 YTLDPHTRWEYCAIKTCADNTMNDTVPLETTETCIGQGEYRGVTNTWNGIPQORWDS 331
 :|||||
 QY 301 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFSTTDPNIRVGYCSQIPNCDMSHGQDC 360
 :|||||
 Db 332 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFSTTDPNIRVGYCSQIPNCDMSHGQDC 391
 :|||||
 QY 361 YRGNGKNYMGNLSQTRSGLTCSMWKDNMEDLHRHIFWEPDASKLNEYCRNPDDDAHGWP 420
 :|||||
 Db 392 YRGNGKNYMGNLSQTRSGLTCSMWKDNMEDLHRHIFWEPDASKLNEYCRNPDDDAHGWP 451
 :|||||
 QY 421 CYTGNPLIPWDYCPISRCGDTTPTIV 447
 :|||||
 Db 452 CYTGNPLIPWDYCPISRCGDTTPTIV 478
 :|||||

RESULT 6

AAR25676
 ID AAR25676 standard; protein; 728 AA.

XX
 AC
 DT 20-JAN-1993 (first entry)

XX
 DE Recombinant human hepatocyte growth factor.
 XX HGF; enhance growth; preparing transgenic animals; hepatic disease;
 KW clinical diagnostic reagent; drug.

XX Homo sapiens.

XX OS

XX JP04183394-A.

XX PN

XX 30-JUN-1992.

XX PD

```

XX 19-NOV-1990; 90JP-0314548.
XX
XX 19-NOV-1990; 90JP-0314548.
XX
XX (NAKA/) NAKAMURA T.
XX (TOYM ) TOYONO KK.
XX
XX WPI; 1992-265591/32.
XX
XX Recombinant human hepatocyte growth factor and DNA encoding it -
XX useful for diagnosis and treatment of hepatic disease and
XX transgenic animal prepn.
XX
XX Disclosure; Page 11; 28pp; Japanese.
XX
XX This sequence represents a recombinant human hepatocyte growth
XX factor. It has physiological activity, and using it enhanced growth
XX of hepatocytes is possible. It is useful as a clinical diagnostic
XX reagent, or a drug for treating hepatic disease.
XX See also AAR25676-92, AAQ26713-27.
XX
XX Sequence 728 AA;

Query Match          99.9%; Score 2601; DB 13; Length 728;
Best Local Similarity 99.8%; Pred. No. 1.4e-165;
Matches 446; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ERKRRNTIHEFKSAKTLTIKIDPALKIKTKKVNADQCANRCKTRNKGLPFTCKAFVFDK 60
DB :|||||
DB 32 QKRRNTIHEFKSAKTLTIKIDPALKIKTKKVNADQCANRCKTRNKGLPFTCKAFVFDK 91
QY 61 ARKQCLWFFPNSMSGVKKFEGHEFDLYENKDYIRNCIIIGKRSYKGTVSIKSGIKCOP 120
DB :|||||
DB 92 ARKQCLWFFPNSMSGVKKFEGHEFDLYENKDYIRNCIIIGKRSYKGTVSIKSGIKCOP 151
QY 121 WSSMIPHEHSFLPSSYRGKDLQENYCRNPRGEGGPGWFTSNPEVRYEVCIDIPOQSEVEC 180
DB :|||||
DB 152 WSSMIPHEHSFLPSSYRGKDLQENYCRNPRGEGGPGWFTSNPEVRYEVCIDIPOQSEVEC 211
QY 181 MTCNGESYRGLMDHTESGKICORWDHQTPIRHKFLPERYPDKGFDNDCRNPDGQPRPWC 240
DB :|||||
DB 212 MTCNGESYRGLMDHTESGKICORWDHQTPIRHKFLPERYPDKGFDNDCRNPDGQPRPWC 271
QY 241 YTLDPHTRWEYCAIKTCADNTMNDTVPLETTECIQOGEGYRGTVNTIWNIGIPQQRWDS 300
DB :|||||
DB 272 YTLDPHTRWEYCAIKTCADNTMNDTVPLETTECIQOGEGYRGTVNTIWNIGIPQQRWDS 331
QY 301 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 360
DB :|||||
DB 332 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 391
QY 361 YRGNGKNYMGNLISQTRSGLTCSMWKDNMEDLHRHIFWEPDASKLNYENCRNPDHAGPW 420
DB :|||||
DB 392 YRGNGKNYMGNLISQTRSGLTCSMWKDNMEDLHRHIFWEPDASKLNYENCRNPDHAGPW 451
QY 421 CYTGNPLIPWDYCPISRCGEGDTTPTIV 447
DB :|||||
DB 452 CYTGNPLIPWDYCPISRCGEGDTTPTIV 478

RESULT 7
AAR40862
ID AAR40862 standard; Protein; 728 AA.
XX
XX AAR40862;
AC
XX
XX 14-MAR-1994 (first entry)
XX
XX Competative inhibitor of HGF.
XX HGF; hepatocyte growth factor; transformation; antagonist;
XX liver disease.

```

```

XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX Peptide 1..31
XX Protein /label= sig_peptide
XX /label= mat_protein
XX Misc-difference 494
XX /note= "site of possible mutation"
XX
XX JP05208998-A.
XX
XX 20-AUG-1993.
XX
XX 25-DEC-1991; 91JP-0357040.
XX
XX 25-DEC-1991; 91JP-0357040.
XX
XX (MITU ) MITSUBISHI KASEI CORP.
XX
XX WPI; 1993-297806/38.
XX
XX N-PSDB; AAQ47832.
XX
XX New protein is competitive inhibitor of hepatocyte growth factor
XX - used in animal model of liver disease
XX
XX Claim 7; Page 8-11; 20pp; Japanese.
XX
XX The gene encodes an inhibitor of HGF, which has a molecular weight
XX of ca. 76000 to 92000 as determined by SDS-PAGE. Human modified
XX (NP) HGF-A and -B fragments were prepared from human HGF cDNA.
XX This was amplified by PCR to give a human NP-HGF DNA fragment that
XX was transformed into E.coli. (see AAQ47833) for a related sequence.
XX
XX Sequence 728 AA;

Query Match          99.9%; Score 2601; DB 14; Length 728;
Best Local Similarity 99.8%; Pred. No. 1.4e-165;
Matches 446; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ERKRRNTIHEFKSAKTLTIKIDPALKIKTKKVNADQCANRCKTRNKGLPFTCKAFVFDK 60
DB :|||||
DB 32 QKRRNTIHEFKSAKTLTIKIDPALKIKTKKVNADQCANRCKTRNKGLPFTCKAFVFDK 91
QY 61 ARKQCLWFFPNSMSGVKKFEGHEFDLYENKDYIRNCIIIGKRSYKGTVSIKSGIKCOP 120
DB :|||||
DB 92 ARKQCLWFFPNSMSGVKKFEGHEFDLYENKDYIRNCIIIGKRSYKGTVSIKSGIKCOP 151
QY 121 WSSMIPHEHSFLPSSYRGKDLQENYCRNPRGEGGPGWFTSNPEVRYEVCIDIPOQSEVEC 180
DB :|||||
DB 152 WSSMIPHEHSFLPSSYRGKDLQENYCRNPRGEGGPGWFTSNPEVRYEVCIDIPOQSEVEC 211
QY 181 MTCNGESYRGLMDHTESGKICORWDHQTPIRHKFLPERYPDKGFDNDCRNPDGQPRPWC 240
DB :|||||
DB 212 MTCNGESYRGLMDHTESGKICORWDHQTPIRHKFLPERYPDKGFDNDCRNPDGQPRPWC 271
QY 241 YTLDPHTRWEYCAIKTCADNTMNDTVPLETTECIQOGEGYRGTVNTIWNIGIPQQRWDS 300
DB :|||||
DB 272 YTLDPHTRWEYCAIKTCADNTMNDTVPLETTECIQOGEGYRGTVNTIWNIGIPQQRWDS 331
QY 301 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 360
DB :|||||
DB 332 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 391
QY 361 YRGNGKNYMGNLISQTRSGLTCSMWKDNMEDLHRHIFWEPDASKLNYENCRNPDHAGPW 420
DB :|||||
DB 392 YRGNGKNYMGNLISQTRSGLTCSMWKDNMEDLHRHIFWEPDASKLNYENCRNPDHAGPW 451
QY 421 CYTGNPLIPWDYCPISRCGEGDTTPTIV 447
DB :|||||
DB 452 CYTGNPLIPWDYCPISRCGEGDTTPTIV 478

```

QY 241 YTLDPHTRWEYCAIKTCADNTMNDTDVPLETTECIQOGEGYRGTYNTIWNIGPCQRWDS 300
DB 272 YTLDPHTRWEYCAIKTCADNTMNDTDVPLETTECIQOGEGYRGTYNTIWNIGPCQRWDS 331
QY 301 QYPHEHMTPEPFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 360
DB 332 QYPHEHMTPEPFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 391
QY 361 YRGNKKNYMGNLSTQSRGLTCSMWKDNMEDLHRRHIFWEPDASKNLNENYCRNPDDAHGPW 420
DB 392 YRGNKKNYMGNLSTQSRGLTCSMWKDNMEDLHRRHIFWEPDASKNLNENYCRNPDDAHGPW 451
QY 421 CYTGNPLIPWDYCPISRCGEGDTTPTIV 447
DB 452 CYTGNPLIPWDYCPISRCGEGDTTPTIV 478

RESULT 9
AAR42062
ID AAR42062 standard; Protein; 728 AA.
XX
AC AAR42062;
XX
DT 15-NOV-1993 (first entry)
XX
DE Vascular endothelial cell growth enhancer protein.
XX
KW Enhance; growth; vascular endothelial cell; human; tumour; cell line;
KW HUOCA-II; HUOCA-III; blood vessel; wounds; burns; decubitus;
XX post-operative tissue damage; drug; cardiac angiopathy.
OS Homo sapiens.
XX
PN EP550296-A.
XX
PD 07-JUL-1993.
XX
PF 27-NOV-1992; 92EP-0403199.
XX
PR 28-NOV-1991; 91JP-0337999.
XX
PA (TERU) TERUMO CORP.
XX
PI Adachi M, Harada K, Hirahara I, Sudo T;
XX
DR WPI; 1993-215669/27.
DR N-PSDB; AAQ45702.
XX
PT Vascular endothelial cell growth factor protein - used for
PT promoting angiogenesis in the treatment of cardiac angiopathy,
PT wounds, burn injuries, postoperative tissue damage etc.
XX
PS Claim 6; Page 18-21; 44pp; English.
XX
CC This sequence represents a single chain protein which selectively
CC enhances the growth of vascular endothelial cells. This protein
CC was produced by the human tumour cell line HUOCA-II or HUOCA-III.
CC This protein enhances the formation of new blood vessels and may be
CC used to enhance healing of wounds, burns injuries, decubitus or
CC post-operative tissue damage. It may also be used as a drug for
CC cardiac angiopathy.
XX
SQ Sequence 728 AA;

Query Match 99.9%; Score 2601; DB 14; Length 728;
Best Local Similarity 99.8%; Pred. No. 1.4e-165;
Matches 446; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ERKRRNTIHEFKSAKTLIKIDPALKIKTKYNTADQCANRCTRNGKLPFTCKAFVFDK 60
DB 32 ORKRRNTIHEFKSAKTLIKIDPALKIKTKYNTADQCANRCTRNGKLPFTCKAFVFDK 91

RESULT 8
AAR40863
ID AAR40863 standard; Protein; 728 AA.
XX
AC AAR40863;
XX
DT 14-MAR-1994 (first entry)
XX
DE Competitive inhibitor of HGF.
XX
KW HGF; hepatocyte growth factor; transformation; antagonist;
XX liver disease.
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Peptide 1..31
FT Protein 32..728 /label= sig_peptide
FT /label= mat_protein
FT Misc-difference 489 /note= "site of possible mutation"
FT Misc-difference 491 /note= "site of possible mutation"
FT Misc-difference 494 /note= "site of possible mutation"
XX
XX JP05208998-A.
XX
XX 20-AUG-1993.
XX
XX 25-DEC-1991; 91JP-0357040.
XX
XX 25-DEC-1991; 91JP-0357040.
XX
XX (MITU) MITSUBISHI KASEI CORP.
XX
XX WPI; 1993-297806/38.
XX
XX N-PSDB; AAQ47833.
XX
XX New protein is competitive inhibitor of hepatocyte growth factor
XX - used in animal model of liver disease
XX
XX Claim 6; Page 11-14; 20pp; Japanese.
XX
XX The gene encodes an inhibitor of HGF, which has a molecular weight
XX of ca. 76000 to 92000 as determined by SDS-PAGE. Human modified
XX (NP) HGF-A and -B fragments were prepared from human HGF cDNA.
XX This was amplified by PCR to give a human NP-HGF DNA fragment that
XX was transformed into E.coli. (see AAQ47833) for a related sequence.
XX
XX Sequence 728 AA;

Query Match 99.9%; Score 2601; DB 14; Length 728;
Best Local Similarity 99.8%; Pred. No. 1.4e-165;
Matches 446; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ERKRRNTIHEFKSAKTLIKIDPALKIKTKYNTADQCANRCTRNGKLPFTCKAFVFDK 60
DB 32 ORKRRNTIHEFKSAKTLIKIDPALKIKTKYNTADQCANRCTRNGKLPFTCKAFVFDK 91

QY 61 ARKQCLWFPPNFMSSGVKKEGHEFDLYENKDYIRNCLIGKRSYKGTSTKSGIKQCP 120
DB 92 ARKQCLWFPPNFMSSGVKKEGHEFDLYENKDYIRNCLIGKRSYKGTSTKSGIKQCP 151

QY 121 WSSMIPHEHSFPLSSYRGKDLOENYCRNPRGEGPGWFTSNPEVRYEVCDIPQCSVEEC 180
DB 152 WSSMIPHEHSFPLSSYRGKDLOENYCRNPRGEGPGWFTSNPEVRYEVCDIPQCSVEEC 211

QY 181 MTCNGESYRGLMDHTESGKICQWRDHQTPHRRKFLPERYPDKGFDNDCYCRNPQGPWC 240
DB 212 MTCNGESYRGLMDHTESGKICQWRDHQTPHRRKFLPERYPDKGFDNDCYCRNPQGPWC 271

Qy 61 ARKQCLWPPFNMSGSGVKKERGFHEFDLYENKDYIRNCIIIGKRSYKGVTSITKSGIKQCP 120
Db 92 ARKQCLWPPFNMSGSGVKKERGFHEFDLYENKDYIRNCIIIGKRSYKGVTSITKSGIKQCP 151
Qy 121 WSSMIPHEHSFLPSSYRGKDLQENYCRNPRGEGGPGWFTSNPEVRYEVCIDIPOCSEVEC 180
Db 152 WSSMIPHEHSFLPSSYRGKDLQENYCRNPRGEGGPGWFTSNPEVRYEVCIDIPOCSEVEC 211
Qy 181 MTCNGESYRGLMDHTESGKICQWRDQHTPHRHKFLPERYPDKGFDNNYCRNPDGQPRWC 240
Db 212 MTCNGESYRGLMDHTESGKICQWRDQHTPHRHKFLPERYPDKGFDNNYCRNPDGQPRWC 271
Qy 241 YLDPHTRWEYCAIKTCADNTMNDTVPLETTECIQOGEGYRGTVNTIWNIGPCQRWDS 300
Db 272 YLDPHTRWEYCAIKTCADNTMNDTVPLETTECIQOGEGYRGTVNTIWNIGPCQRWDS 331
Qy 301 OYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 360
Db 332 OYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 391
Qy 361 YRGNGKNTMGNSLQSTRSLGTCSMWDMKMDLHRHIFWEPDASKLNENYCRNPDDDAHGWP 420
Db 392 YRGNGKNTMGNSLQSTRSLGTCSMWDMKMDLHRHIFWEPDASKLNENYCRNPDDDAHGWP 451
Qy 421 CYTGNPLIPWDYCPISRCGEGDTTPTIV 447
Db 452 CYTGNPLIPWDYCPISRCGEGDTTPTIV 478

RESULT 10

AAW00338
ID AAW00338 standard; protein; 728 AA.
AC AAW00338;
XX 27-JUN-1997 (first entry)
XX Human hepatic parenchymal cell growth factor.
XX Human; hepatic; parenchymal cell; growth factor; HGF; treatment;
KW prevention; ischaemia; ischaemic; disease; reperfusion; disorder;
KW blood; liver; transplant; acute; failure; ischemia; ischemic;
KW organ.
XX Homo sapiens.
OS WO9632960-A1.
PN 24-OCT-1996.
XX 19-APR-1996; 96WO-JP01065.
XX 21-APR-1995; 95JP-0096994.
XX (MITU) MITSUBISHI CHEM CORP.
XX Gamba M, Yonehana T;
XX WPI; 1996-485555/48.
XX Hepatic parenchymal cell growth factors - used as remedy and/or
PT preventive medicine for ischaemic diseases
XX Claim 9; Pages 7-10; 16pp; Japanese.
XX The present sequence is the human hepatic parenchymal cell growth
CC factor (HGF), which has a molecular weight of 76-92 kD estimated by
CC SDS-PAGE, parenchymal cell growth activity, which is lost following
CC heat treatment at 80 degrees C for 10 minutes and strong affinity
CC towards heparin. It may be used to treat and prevent ischaemic
CC diseases, e.g. blood reperfusion disorder, ischaemic reperfusion
CC disorders in liver transplant, acute liver failure and organ
CC disorder during organ transplant. For adults, it is administered at

CC a level of 1 microg to 10 mg/kg/day, preferably 10-1000
CC microg/kg/day.
CC A pig liver derived cultured epithelial cell line was cultured in
CC DMEM, cow embryo serum and N-2-hydroxyethylpiperazine ethane
CC sulphonic acid to confluence. The medium was then changed to serum
CC and glucose free DMEM, and the oxygen concentration reduced to less
CC than 2%. The cells were cultured for 6 hours, and then for another
CC hour at 95% air/5% carbon dioxide (reoxygenation). HGF was added
CC just before hypoxiation or reoxygenation. HGF suppressed free
CC lactic acid dehydrogenase (LDH) by 49.6% when added before
CC hypoxiation, and by 60.9% when added before reoxygenation, at 50
CC ng/ml (free LDL was 100% in the absence of HGF).
XX
SQ Sequence 728 AA;
Query Match 99.9%; Score 2601; DB 17; Length 728;
Best Local Similarity 99.8%; Pred. No. 1.4e-165;
Matches 446; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ERKRNTIHEFKSAKTTLIKIDPALKIKTKKVNADQCANRCKTRNKGLPFTCKAFVFDK 60
Db 32 QKRKNTIHEFKSAKTTLIKIDPALKIKTKKVNADQCANRCKTRNKGLPFTCKAFVFDK 91
Qy 61 ARKQCLWPPFNMSGSGVKKERGFHEFDLYENKDYIRNCIIIGKRSYKGVTSITKSGIKQCP 120
Db 92 ARKQCLWPPFNMSGSGVKKERGFHEFDLYENKDYIRNCIIIGKRSYKGVTSITKSGIKQCP 151
Qy 121 WSSMIPHEHSFLPSSYRGKDLQENYCRNPRGEGGPGWFTSNPEVRYEVCIDIPOCSEVEC 180
Db 152 WSSMIPHEHSFLPSSYRGKDLQENYCRNPRGEGGPGWFTSNPEVRYEVCIDIPOCSEVEC 211
Qy 181 MTCNGESYRGLMDHTESGKICQWRDQHTPHRHKFLPERYPDKGFDNNYCRNPDGQPRWC 240
Db 212 MTCNGESYRGLMDHTESGKICQWRDQHTPHRHKFLPERYPDKGFDNNYCRNPDGQPRWC 271
Qy 241 YLDPHTRWEYCAIKTCADNTMNDTVPLETTECIQOGEGYRGTVNTIWNIGPCQRWDS 300
Db 272 YLDPHTRWEYCAIKTCADNTMNDTVPLETTECIQOGEGYRGTVNTIWNIGPCQRWDS 331
Qy 301 OYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 360
Db 332 OYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 391
Qy 361 YRGNGKNTMGNSLQSTRSLGTCSMWDMKMDLHRHIFWEPDASKLNENYCRNPDDDAHGWP 420
Db 392 YRGNGKNTMGNSLQSTRSLGTCSMWDMKMDLHRHIFWEPDASKLNENYCRNPDDDAHGWP 451
Qy 421 CYTGNPLIPWDYCPISRCGEGDTTPTIV 447
Db 452 CYTGNPLIPWDYCPISRCGEGDTTPTIV 478

RESULT 11

AAW00340
ID AAW00340 standard; protein; 728 AA.
XX
AC AAW00340;
XX 09-DEC-1996 (first entry)
XX Wild type hepatocyte growth factor.
DE Human; hepatocyte growth factor; HGF; huHGF; serum; proteolytic cleavage;
XX pro-hormone; beta subunit; alpha subunit; kringle domain; prothrombin;
KW plasminogen; catalytic domain; serine protease; HGF variant;
KW HGF receptor; malignancy; chronic HGF receptor activation.
XX Homo sapiens.
OS
XX Key Location/Qualifiers
FH Cleavage-site 494..495
FT /label= Proteolytic_cleavage_site
FT /note= "generates alpha and beta subunits"

FT Disulfide-bond 487..604
FT Peptide 1..54
FT /note= "prepro-sequence"
FT Peptide 1..31
FT /note= "Hydrophobic signal peptide"
FT Domain 128..206
FT /label= Kringle_1_domain
FT Domain 211..288
FT /label= Kringle_2_domain
FT Domain 303..383
FT /label= Kringle_3_domain
FT Domain 391..464
FT /label= Kringle_4_domain
FT Modified-site 294
FT /label= N-linked-glycosylation_site
FT Modified-site 402
FT /label= N-linked-glycosylation_site
FT Modified-site 566
FT /label= N-linked-glycosylation_site
FT Modified-site 653
FT /label= N-linked-glycosylation_site
FT Misc-difference 494
FT /note= "Position of opt. substitution, esp. Glu, Asp or Asn"
FT Misc-difference 495
FT /note= "Position of opt. substitution, esp. Tyr or Phe"
FT Misc-difference 534
FT /note= "Position of opt. substitution"
FT Misc-difference 673
FT /note= "Position of opt. substitution"
FT Misc-difference 692
FT /note= "Position of opt. substitution"
FT FT
XX US5547856-A.
XX 20-AUG-1996.
XX 18-MAY-1992; 92US-0884811.
XX 13-JUL-1993; 93US-0087783.
XX 18-MAY-1992; 92US-0884811.
XX 18-MAY-1992; 92US-0885971.
XX (GETH) GENENTECH INC.
XX Godowski PJ, Lokker NA, Mark MR;
XX WPI; 1996-392634/39.
XX New hepatocyte growth factor variants - are resistant to in vivo
XX proteolytic cleavage into a 2-chain form, useful as HGF antagonists
XX Disclosure: Column 51-56; 39pp; English.
XX This sequence represents the wild type full length sequence of human
XX hepatocyte growth factor (HGF). HGF is isolated from human serum
XX and is a disulphide linked heterodimer derived by proteolytic cleavage
XX of this pro-hormone form between residues 494 and 495. This generates a
XX molecule composed of an alpha-subunit of 440 amino acids (mol. wt. 69
XX kD) and a beta-subunit of 234 amino acids (mol. wt. 34 kD). The alpha
XX and beta subunits are encoded by a single open reading frame. The alpha
XX subunit contains four kringle domains based on their homology to
XX kringle-like domains in other proteins, e.g. prothrombin, plasminogen.
XX The beta subunit shows high homology to the catalytic domain of serine
XX proteases. However two of the three residues which form the catalytic
XX triad of serine proteases are not conserved in HGF. Therefore, the
XX precise function of the beta chain remains unknown. The invention
XX includes HGF variants which retain HGF receptor binding activity without
XX having the biological activity of wild-type HGF. They can be used for
XX the treatment of pathological conditions associated with the activation
XX of a HGF receptor such as malignancies associated with chronic HGF
XX receptor activation.

SQ	Sequence	728 AA;
Query Match	99.9%;	Score 2601; DB 17; Length 728;
Best Local Similarity	99.8%;	Pred. No. 1.4e-165;
Matches 446; Conservative	1;	Mismatches 0; Indels 0; Gaps 0;
QY	1	ERRRRNTIHEFKSAKTTLLIKIDPALKIKTKKVTADQCANRCKTRNKGLPFTCKAFVFK 60
DB	32	QRRRRNTIHEFKSAKTTLLIKIDPALKIKTKKVTADQCANRCKTRNKGLPFTCKAFVFK 91
QY	61	ARKOCLWFPNSSSGVKKEFGHEFDLYENKDYIRNCIIIGKGRSYKGTYSITKSGIKCOP 120
DB	92	ARKOCLWFPNSSSGVKKEFGHEFDLYENKDYIRNCIIIGKGRSYKGTYSITKSGIKCOP 151
QY	121	WSSMTPHEHSFLPSSYRGKDLQENYCRNPRGEGGPMCFITSNPEVRYEVCIDIQCSEVEC 180
DB	152	WSSMTPHEHSFLPSSYRGKDLQENYCRNPRGEGGPMCFITSNPEVRYEVCIDIQCSEVEC 211
QY	181	MTNCGESYRGIMDHTEGKICQWRDHOTPHRHKFLPERYPDKGFDNNYCRNPDGQPRWC 240
DB	212	MTNCGESYRGIMDHTEGKICQWRDHOTPHRHKFLPERYPDKGFDNNYCRNPDGQPRWC 271
QY	241	YTLDPHTRWEYCAIKTCADNTMNDTDVPLETTECICQGGEGYRGTNTIWNIGIPCORWDS 300
DB	272	YTLDPHTRWEYCAIKTCADNTMNDTDVPLETTECICQGGEGYRGTNTIWNIGIPCORWDS 331
QY	301	OYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 360
DB	332	OYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 391
QY	361	YRGNGKNTYMGNLSOTRSLTCSMDKKNMEDLHRHIFWEPDASKLNENYCRNPDGDAHGPW 420
DB	392	YRGNGKNTYMGNLSOTRSLTCSMDKKNMEDLHRHIFWEPDASKLNENYCRNPDGDAHGPW 451
QY	421	CYTGNPLIPWDYCPISRCGEGDTTPTIV 447
DB	452	CYTGNPLIPWDYCPISRCGEGDTTPTIV 478
RESULT 12		
AAR87522		
ID	AAR87522 standard; protein; 728 AA.	
XX		
AC	AAR87522;	
XX		
DT	21-JUN-1996 (first entry)	
XX		
DE	Mutant hepatocyte growth factor #1 with changes at residues 491-495.	
XX		
KW	Hepatocyte growth factor; protease; cleavage recognition site; Factor Xa;	
KW	injury; blood vessel; point mutation.	
OS	Synthetic.	
XX		
XX	Key Location/Qualifiers	
FT	Misc-difference 491..495	
FT	/note= "change from wild type sequence:	
FT	Lys-Glu-Leu-Arg-Val to Ile-Glu-Gly-Arg-Thr"	
XX		
PN	JP07304796-A.	
XX		
PD	21-NOV-1995.	
XX		
PF	07-MAY-1994; 94JP-0117506.	
XX		
PR	07-MAY-1994; 94JP-0117506.	
XX		
PA	(TERU) TERUMO CORP.	
XX		
DR	WPI; 1996-035890/04.	
XX		
PT	Novel derivative of hepatocyte growth factor - comprises a mutation	
PT	at positions 492-494, for use in treating injured blood vessels	

```

XX PS Disclosure: Page -: 8pp; Japanese.
XX
CC The amino acid of the hepatocyte growth factor (HGF) generated to
CC contain amino acid changes at residues 492-494. (Note - this sequence is
CC not given in the specification but is based on the HGF sequence given in
CC AAR20005). The generated mutant contains the sequence
CC Ile-Glu-Arg-Thr, which replaces the wild type sequence
CC Lys-Glu-Leu-Arg-Val. The HGF is translated as single chain protein which
CC is activated to a two chain protein by protease cleavage between residues
CC 494-5, producing the alpha and beta chains. The novel sequence alters the
CC protease recognition site, putatively introducing a factor Xa recognition
CC site. The novel HGF deriv. can be activated specifically at the site of
CC injury in a blood vessel. (See AAR87522-5 for different HGF derivs.). The
CC novel HGF derivs. were generated by point mutations using the
CC oligonucleotides AAT06762-5.
XX
SQ Sequence 728 AA;

Query Match 99.9%; Score 2601; DB 17; Length 728;
Best Local Similarity 99.8%; Pred. No. 1.4e-165;
Matches 446; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ERKRNTHEFKSAKTLIKIDPALKIKTKKVNADQCANRCKTRNKGLPFTCKAFVFDK 60
Db :|||||
Qy 32 QKRKNTHEFKSAKTLIKIDPALKIKTKKVNADQCANRCKTRNKGLPFTCKAFVFDK 91
Db :|||||
Qy 61 ARKQCLWFPFNSMSGVKKEFGHEFDLYENKDYIRNCIIIGKGRSYKGTVSIITKSGIKQCP 120
Db :|||||
Qy 92 ARKQCLWFPFNSMSGVKKEFGHEFDLYENKDYIRNCIIIGKGRSYKGTVSIITKSGIKQCP 151
Db :|||||
Qy 121 WSSMTPHEHSFLPSSYRGKDLOENYCRNPRGEGGPGWCFSTSNPEVRYEVCIDIPOCSEVEC 180
Db :|||||
Qy 152 WSSMTPHEHSFLPSSYRGKDLOENYCRNPRGEGGPGWCFSTSNPEVRYEVCIDIPOCSEVEC 211
Db :|||||
Qy 181 MTCNGESYRGLMDHTESGKICORWDHQPTRHKFLPERYPDKGFDNYCRNPDGQPRPWC 240
Db :|||||
Qy 212 MTCNGESYRGLMDHTESGKICORWDHQPTRHKFLPERYPDKGFDNYCRNPDGQPRPWC 271
Db :|||||
Qy 241 YLDPHTRWEYCAIKTCADNTMNDTDVPLETTECIQOGGEGYRGVTNTWNGIPCORWDS 300
Db :|||||
Qy 272 YLDPHTRWEYCAIKTCADNTMNDTDVPLETTECIQOGGEGYRGVTNTWNGIPCORWDS 331
Db :|||||
Qy 301 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 360
Db :|||||
Qy 332 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 391
Db :|||||
Qy 361 YRGNCKNTMGNLSQTRSGLTCSMWKDNMEDLHRHIFWEPDASKLNEYCRNPDGDDAHGPW 420
Db :|||||
Qy 392 YRGNCKNTMGNLSQTRSGLTCSMWKDNMEDLHRHIFWEPDASKLNEYCRNPDGDDAHGPW 451
Db :|||||
Qy 421 CYTGNPLIPWDYCPISRCGEGDTTPTIV 447
Db :|||||
Qy 452 CYTGNPLIPWDYCPISRCGEGDTTPTIV 478
Db :|||||

RESULT 13
AAR87523
ID AAR87523 standard; protein; 728 AA.
XX
AC AAR87523;
XX
DT 21-JUN-1996 (first entry)
XX
DE Mutant hepatocyte growth factor #2 with changes at residues 491-495.
XX
KW Hepatocyte growth factor; protease; cleavage recognition site; Factor Xa;
KW injury; blood vessel; point mutation.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 491..495

```

```

FT FT /note= "change from wild type sequence:
XX XX Lys-Glu-Leu-Arg-Val to Lys-Gln-Arg-Ile"
PN JP07304796-A.
XX
PD 21-NOV-1995.
XX
PF 07-MAY-1994; 94JP-0117506.
XX
PR 07-MAY-1994; 94JP-0117506.
XX
PA (TERU ) TERUMO CORP.
XX
DR WPI; 1996-035890/04.
XX
PT Novel derivative of hepatocyte growth factor - comprises a mutation
PT at positions 492-494, for use in treating injured blood vessels
XX
PS Disclosure: Page -: 8pp; Japanese.
XX
CC The amino acid of the hepatocyte growth factor (HGF) generated to
CC contain amino acid changes at residues 492-494. (Note - this sequence is
CC not given in the specification but is based on the HGF sequence given in
CC AAR20005). The generated mutant contains the sequence
CC Lys-Gln-Arg-Val, which replaces the wild type sequence
CC Lys-Glu-Leu-Arg-Val. The HGF is translated as single chain protein which
CC is activated to a two chain protein by protease cleavage between residues
CC 494-5, producing the alpha and beta chains. The novel sequence alters the
CC protease recognition site, putatively introducing a Factor Xa recognition
CC site. The novel HGF deriv. can be activated specifically at the site of
CC injury in a blood vessel. (See AAR87522-5 for different HGF derivs.). The
CC novel HGF derivs. were generated by point mutations using the
CC oligonucleotides AAT06762-5.
XX
SQ Sequence 728 AA;

Query Match 99.9%; Score 2601; DB 17; Length 728;
Best Local Similarity 99.8%; Pred. No. 1.4e-165;
Matches 446; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ERKRNTHEFKSAKTLIKIDPALKIKTKKVNADQCANRCKTRNKGLPFTCKAFVFDK 60
Db :|||||
Qy 32 QKRKNTHEFKSAKTLIKIDPALKIKTKKVNADQCANRCKTRNKGLPFTCKAFVFDK 91
Db :|||||
Qy 61 ARKQCLWFPFNSMSGVKKEFGHEFDLYENKDYIRNCIIIGKGRSYKGTVSIITKSGIKQCP 120
Db :|||||
Qy 92 ARKQCLWFPFNSMSGVKKEFGHEFDLYENKDYIRNCIIIGKGRSYKGTVSIITKSGIKQCP 151
Db :|||||
Qy 121 WSSMTPHEHSFLPSSYRGKDLOENYCRNPRGEGGPGWCFSTSNPEVRYEVCIDIPOCSEVEC 180
Db :|||||
Qy 152 WSSMTPHEHSFLPSSYRGKDLOENYCRNPRGEGGPGWCFSTSNPEVRYEVCIDIPOCSEVEC 211
Db :|||||
Qy 181 MTCNGESYRGLMDHTESGKICORWDHQPTRHKFLPERYPDKGFDNYCRNPDGQPRPWC 240
Db :|||||
Qy 212 MTCNGESYRGLMDHTESGKICORWDHQPTRHKFLPERYPDKGFDNYCRNPDGQPRPWC 271
Db :|||||
Qy 241 YLDPHTRWEYCAIKTCADNTMNDTDVPLETTECIQOGGEGYRGVTNTWNGIPCORWDS 300
Db :|||||
Qy 272 YLDPHTRWEYCAIKTCADNTMNDTDVPLETTECIQOGGEGYRGVTNTWNGIPCORWDS 331
Db :|||||
Qy 301 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 360
Db :|||||
Qy 332 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 391
Db :|||||
Qy 361 YRGNCKNTMGNLSQTRSGLTCSMWKDNMEDLHRHIFWEPDASKLNEYCRNPDGDDAHGPW 420
Db :|||||
Qy 392 YRGNCKNTMGNLSQTRSGLTCSMWKDNMEDLHRHIFWEPDASKLNEYCRNPDGDDAHGPW 451
Db :|||||
Qy 421 CYTGNPLIPWDYCPISRCGEGDTTPTIV 447
Db :|||||
Qy 452 CYTGNPLIPWDYCPISRCGEGDTTPTIV 478
Db :|||||

```


QY 1 ERKRNTIHEFKSAKTTLIKIDPALKIKTKKVNADQCANRCTRNKGLPFTCKAFVFDK 60
Db :|||||
QY 32 QKRRNTIHEFKSAKTTLIKIDPALKIKTKKVNADQCANRCTRNKGLPFTCKAFVFDK 91
Db :|||||
QY 61 ARKQCLWFPFNSMSGVKEFGHEFDLYENKDYIRNCIIGKRSYKGTVSIKSGIKCOP 120
Db :|||||
QY 92 ARKQCLWFPFNSMSGVKEFGHEFDLYENKDYIRNCIIGKRSYKGTVSIKSGIKCOP 151
Db :|||||
QY 121 WSSMIPHEHSFLPSSYRGKDLQENYCRNPRGEGGPGWCFSTSNPEVRYEVCIDIPOCSEVBC 180
Db :|||||
QY 152 WSSMIPHEHSFLPSSYRGKDLQENYCRNPRGEGGPGWCFSTSNPEVRYEVCIDIPOCSEVBC 211
Db :|||||
QY 181 MTCNGESYRGLMDHTESGKICQWHDQTPHRHKFLPERYPDKGFDNYCRNPDGQPRPMC 240
Db :|||||
QY 212 MTCNGESYRGLMDHTESGKICQWHDQTPHRHKFLPERYPDKGFDNYCRNPDGQPRPMC 271
Db :|||||
QY 241 YTLDPHTRWEXCAIKTCADNTMNDTDVPLETTECIQOGGEGYRGTVNTIWNIGPCQWDS 300
Db :|||||
QY 272 YTLDPHTRWEXCAIKTCADNTMNDTDVPLETTECIQOGGEGYRGTVNTIWNIGPCQWDS 331
Db :|||||
QY 301 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 360
Db :|||||
QY 332 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 391
Db :|||||
QY 361 YRGNGKNTYMGNLSTQTRSGLTCSMDKNMEDLHRHIFWEPDASKLNEYCRNPDGDAHGPW 420
Db :|||||
QY 392 YRGNGKNTYMGNLSTQTRSGLTCSMDKNMEDLHRHIFWEPDASKLNEYCRNPDGDAHGPW 451
Db :|||||
QY 421 CYTGNPLIPWDYCPISRCGEGDTTPTIV 447
Db :|||||
QY 452 CYTGNPLIPWDYCPISRCGEGDTTPTIV 478
Db :|||||

Search completed: April 25, 2003, 15:50:29
Job time : 83 secs

THIS PAGE BLANK (uspto)

GenCore version 5.1.4_p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 25, 2003, 15:40:22 ; Search time 18 seconds
(without alignments)
730.669 Million cell updates/sec

Title: us-09-674-377b-1

Perfect score: 2604

Sequence: 1 ERKRRNTIHEFKSAKTLI.....IPWDYCPISRCGDTTPTIV 447

Scoring table: BLOSUM62

Gapop 10.0 ..Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_AA.*

- 1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
- 2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
- 3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
- 4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
- 5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep.*
- 6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2601	99.9	697	2	US-08-460-890A-50
2	2601	99.9	697	3	US-08-167-641C-50
3	2601	99.9	697	4	US-08-460-971A-50
4	2601	99.9	697	4	US-08-462-040-50
5	2601	99.9	728	1	US-07-815-333A-2
6	2601	99.9	728	1	US-08-087-783A-22
7	2563.5	98.4	723	1	US-07-838-410-1
8	2563.5	98.4	723	1	US-08-290-937B-1
9	2563.5	98.4	723	1	US-08-290-937B-2
10	2563.5	98.4	723	1	US-08-404-643-1
11	2563.5	98.4	723	4	US-08-194-326-1
12	2560.5	98.3	723	4	US-08-030-410-3
13	2557.5	98.2	723	1	US-08-290-937B-3
14	2541.5	97.6	723	4	US-08-194-326-2
15	2541.5	97.6	723	4	US-08-700-519J-18
16	2539.5	97.5	723	4	US-08-194-326-3
17	2539.5	97.5	723	4	US-08-700-519J-19
18	1180.5	45.3	711	1	US-08-184-012C-8
19	1180.5	45.3	711	1	US-08-334-177-2
20	1180.5	45.3	711	2	US-08-666-082B-1
21	1180.5	45.3	711	2	US-08-766-982-2
22	1180.5	45.3	711	4	US-09-296-219-2
23	1180.5	45.3	711	5	PCT-US95-13830-2
24	1125	43.2	716	2	US-08-766-982-1
25	1125	43.2	716	4	US-09-296-219-1
26	1059	40.7	812	1	US-08-248-629A-1
27	1059	40.7	812	1	US-08-451-932-1

28	1059	40.7	812	1	US-08-452-260-1	Sequence 1, Appli
29	1059	40.7	812	1	US-08-326-785-1	Sequence 1, Appli
30	1059	40.7	812	2	US-08-612-788-1	Sequence 1, Appli
31	1059	40.7	812	2	US-08-605-598B-1	Sequence 1, Appli
32	1059	40.7	812	2	US-08-429-743-1	Sequence 1, Appli
33	1059	40.7	812	2	US-08-866-735-1	Sequence 1, Appli
34	1059	40.7	812	3	US-09-066-028-1	Sequence 1, Appli
35	1059	40.7	812	5	PCT-US95-05107-1	Sequence 1, Appli
36	1045.5	40.1	790	2	US-08-469-486-54	Sequence 54, Appl
37	1045.5	40.1	790	2	US-08-469-658-54	Sequence 54, Appl
38	1042	40.0	791	1	US-08-643-219-1	Sequence 1, Appli
39	1042	40.0	791	2	US-09-131-995-1	Sequence 1, Appli
40	1042	40.0	791	2	US-08-832-087B-1	Sequence 1, Appli
41	1042	40.0	791	3	US-08-851-350-1	Sequence 1, Appli
42	1042	40.0	791	4	US-09-132-154-1	Sequence 1, Appli
43	1042	40.0	810	1	US-07-854-603-2	Sequence 2, Appli
44	1042	40.0	810	1	US-08-147-000B-29	Sequence 29, Appli
45	1042	40.0	810	4	US-09-086-514-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-08-460-890A-50
; Sequence 50, Application US/08460890A
; Patent No. 5994109
; GENERAL INFORMATION:
; APPLICANT: Wco, Savio L.C.
; APPLICANT: Smith, Louis C.
; APPLICANT: Cristiano, Richard J.
; APPLICANT: Gottchalk, Stephen
; TITLE OF INVENTION: NUCLEIC ACID TRANSPORTER SYSTEMS AND
; TITLE OF INVENTION: METHODS OF USE
; NUMBER OF SEQUENCES: 65
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; STREET: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: FastSeq for Windows 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/460.890A
; FILING DATE: June 5, 1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/167,641
; FILING DATE: December 14, 1993
; APPLICATION NUMBER: 07/855,389
; FILING DATE: March 20, 1992
; APPLICATION NUMBER: PCT/US93/02725
; FILING DATE: March 19, 1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Warburg, Richard J.
; REGISTRATION NUMBER: 32,327
; REFERENCE/DOCKET NUMBER: 212/066
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 50:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 697 amino acids
; TYPE: amino acid
; STRANDEDNESS: single

TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-460-890A-50

Query Match 99.9%; Score 2601; DB 2; Length 697;
Best Local Similarity 99.8%; Pred. No. 1.2e-214;
Matches 446; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 ERKRRNTHEFKKSAKTLIKIDPALKIKTKKVTADOCANRCTRNKGLPTCKAFVFDK 60
Db 1 QKRRNTHEFKKSAKTLIKIDPALKIKTKKVTADOCANRCTRNKGLPTCKAFVFDK 60

QY 61 ARKQCLWFPFNSMSSGVKKEFGHEFDLYENKDYIRNCIIIGKRSYKGTVSITKSGIKQOP 120
Db 61 ARKQCLWFPFNSMSSGVKKEFGHEFDLYENKDYIRNCIIIGKRSYKGTVSITKSGIKQOP 120

QY 121 WSSMIPHEHSLPSSYRGKDLQENYCRNPRGEGGFWCFTSNPEVRYEVCDDIPQCEVEEC 180
Db 121 WSSMIPHEHSLPSSYRGKDLQENYCRNPRGEGGFWCFTSNPEVRYEVCDDIPQCEVEEC 180

QY 181 MTCNGESYRGLMDHTESGKICORWDHQPPIRHKFLPERYPDKGFDNCRNPDGQPRPWC 240
Db 181 MTCNGESYRGLMDHTESGKICORWDHQPPIRHKFLPERYPDKGFDNCRNPDGQPRPWC 240

QY 241 YTLDPHTRWEYCAIKTCADNTMNDTVPLETTETCIGQGEGYRGTVNTIWIWGIPQWRDS 300
Db 241 YTLDPHTRWEYCAIKTCADNTMNDTVPLETTETCIGQGEGYRGTVNTIWIWGIPQWRDS 300

QY 301 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 360
Db 301 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 360

QY 361 YRGNKKNYMGNSQTRSGLTCSMDKNMEDLHRHIFWEPDASKLKNENYCRNPDDDAHGWP 420
Db 361 YRGNKKNYMGNSQTRSGLTCSMDKNMEDLHRHIFWEPDASKLKNENYCRNPDDDAHGWP 420

QY 421 CYTGNPLIPWDYCPISRCBGDTTPTIV 447
Db 421 CYTGNPLIPWDYCPISRCBGDTTPTIV 447
```

RESULT 2

US-08-167-641C-50
Sequence 50, Application US/08167641C
Patent No. 6033884

GENERAL INFORMATION:

APPLICANT: Woo, Savio L.C.
APPLICANT: Smith, Louis C.
APPLICANT: Cristiano, Richard J.
APPLICANT: Gottchalk, Stephen
TITLE OF INVENTION: NUCLEIC ACID TRANSPORTER SYSTEMS AND
METHODS OF USE
NUMBER OF SEQUENCES: 65
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071-2066

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
MEDIUM TYPE: storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: FASTSEQ for Windows 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/167,641C
FILING DATE: December 14, 1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/855,389

FILING DATE: March 20, 1992
APPLICATION NUMBER: PCT/US93/02725
FILING DATE: March 19, 1993
ATTORNEY/AGENT INFORMATION:
NAME: Warburg, Richard J.
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 205/012
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510

INFORMATION FOR SEQ. ID NO. 50:

SEQUENCE CHARACTERISTICS:
LENGTH: 697 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-167-641C-50

Query Match 99.9%; Score 2601; DB 3; Length 697;
Best Local Similarity 99.8%; Pred. No. 1.2e-214;
Matches 446; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 ERKRRNTHEFKKSAKTLIKIDPALKIKTKKVTADOCANRCTRNKGLPTCKAFVFDK 60
Db 1 QKRRNTHEFKKSAKTLIKIDPALKIKTKKVTADOCANRCTRNKGLPTCKAFVFDK 60

QY 61 ARKQCLWFPFNSMSSGVKKEFGHEFDLYENKDYIRNCIIIGKRSYKGTVSITKSGIKQOP 120
Db 61 ARKQCLWFPFNSMSSGVKKEFGHEFDLYENKDYIRNCIIIGKRSYKGTVSITKSGIKQOP 120

QY 121 WSSMIPHEHSLPSSYRGKDLQENYCRNPRGEGGFWCFTSNPEVRYEVCDDIPQCEVEEC 180
Db 121 WSSMIPHEHSLPSSYRGKDLQENYCRNPRGEGGFWCFTSNPEVRYEVCDDIPQCEVEEC 180

QY 181 MTCNGESYRGLMDHTESGKICORWDHQPPIRHKFLPERYPDKGFDNCRNPDGQPRPWC 240
Db 181 MTCNGESYRGLMDHTESGKICORWDHQPPIRHKFLPERYPDKGFDNCRNPDGQPRPWC 240

QY 241 YTLDPHTRWEYCAIKTCADNTMNDTVPLETTETCIGQGEGYRGTVNTIWIWGIPQWRDS 300
Db 241 YTLDPHTRWEYCAIKTCADNTMNDTVPLETTETCIGQGEGYRGTVNTIWIWGIPQWRDS 300

QY 301 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 360
Db 301 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 360

QY 361 YRGNKKNYMGNSQTRSGLTCSMDKNMEDLHRHIFWEPDASKLKNENYCRNPDDDAHGWP 420
Db 361 YRGNKKNYMGNSQTRSGLTCSMDKNMEDLHRHIFWEPDASKLKNENYCRNPDDDAHGWP 420

QY 421 CYTGNPLIPWDYCPISRCBGDTTPTIV 447
Db 421 CYTGNPLIPWDYCPISRCBGDTTPTIV 447
```

RESULT 3

US-08-460-971A-50
Sequence 50, Application US/08460971A
Patent No. 6150168

GENERAL INFORMATION:

APPLICANT: Woo, Savio L.C.
APPLICANT: Smith, Louis C.
APPLICANT: Cristiano, Richard J.
APPLICANT: Gottchalk, Stephen
TITLE OF INVENTION: NUCLEIC ACID TRANSPORTER SYSTEMS AND
METHODS OF USE
NUMBER OF SEQUENCES: 65
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071-2066

```

; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: FastSEQ for Windows 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/460.971A
; FILING DATE: June 5, 1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/167.641
; FILING DATE: December 14, 1993
; APPLICATION NUMBER: 07/855.389
; FILING DATE: March 20, 1992
; APPLICATION NUMBER: PCT/US93/02725
; FILING DATE: March 19, 1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Warburg, Richard J.
; REGISTRATION NUMBER: 32,327
; REFERENCE/DOCKET NUMBER: 212/063
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 50:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 697 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-460-971A-50

```

Query Match 99.9%; Score 2601; DB 4; Length 697;
Best Local Similarity 99.8%; Pred. No. 1.2e-214;
Matches 446; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 EKRRNTTHERPKSAKTLIKIDPALKIKTKKVNADQCANCRNKGLPFTCKAFVFDK 60
DB :|||||
DB 1 QKRRNTTHERPKSAKTLIKIDPALKIKTKKVNADQCANCRNKGLPFTCKAFVFDK 60
QY 61 ARKQCLWFPFNSMSSGVKKEFGHEFDLYENKDYIRNCIIIGKRSYKGTVITKSGIKCOP 120
DB :|||||
DB 61 ARKQCLWFPFNSMSSGVKKEFGHEFDLYENKDYIRNCIIIGKRSYKGTVITKSGIKCOP 120
QY 121 WSSMIPHEHSFLPSSYRGKDLQENYCRNPRGEGGPGWCFTSNPEVRYEYCDIPQCSEVEC 180
DB :|||||
DB 121 WSSMIPHEHSFLPSSYRGKDLQENYCRNPRGEGGPGWCFTSNPEVRYEYCDIPQCSEVEC 180
QY 181 MTCNGESYRGLMDHTESGKICQWHDQTPHRRHKFLPERYPDKGFDNCRNPDGQPRPWC 240
DB :|||||
DB 181 MTCNGESYRGLMDHTESGKICQWHDQTPHRRHKFLPERYPDKGFDNCRNPDGQPRPWC 240
QY 241 YTLDPHTTWEYCAIKTCADNTMNDTVDPLETTECIGOGEGYRGVTNTLWNGIPQORWDS 300
DB :|||||
DB 241 YTLDPHTTWEYCAIKTCADNTMNDTVDPLETTECIGOGEGYRGVTNTLWNGIPQORWDS 300
QY 301 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 360
DB :|||||
DB 301 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 360
QY 361 YRGCKNYMGNLSQTRSGLTCSMWKDNKEDLHRHFWEPDASKLNEYCRNPDHAGHPW 420
DB :|||||
DB 361 YRGCKNYMGNLSQTRSGLTCSMWKDNKEDLHRHFWEPDASKLNEYCRNPDHAGHPW 420
QY 421 CYTGNPLIPWDYCPISRCGDDTPTIV 447
DB :|||||
DB 421 CYTGNPLIPWDYCPISRCGDDTPTIV 447

```

RESULT 4

```

US-08-462-040-50
; Sequence 50, Application US/08462040
; Patent No. 6177554
; GENERAL INFORMATION:
; APPLICANT: Woo, Savio L.C.
; APPLICANT: Smith, Louis C.
; APPLICANT: Cristiano, Richard J.
; APPLICANT: Gottchalk, Stephen
; TITLE OF INVENTION: NUCLEIC ACID TRANSPORTER SYSTEMS AND
; METHODS OF USE
; NUMBER OF SEQUENCES: 65
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; STREET: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: FastSEQ for Windows 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/462.040
; FILING DATE: June 5, 1995
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/167.641
; FILING DATE: December 14, 1993
; APPLICATION NUMBER: 07/855.389
; FILING DATE: March 20, 1992
; APPLICATION NUMBER: PCT/US93/02725
; FILING DATE: March 19, 1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Warburg, Richard J.
; REGISTRATION NUMBER: 32,327
; REFERENCE/DOCKET NUMBER: 212/078
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 50:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 697 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-462-040-50

```

Query Match 99.9%; Score 2601; DB 4; Length 697;
Best Local Similarity 99.8%; Pred. No. 1.2e-214;
Matches 446; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 EKRRNTTHERPKSAKTLIKIDPALKIKTKKVNADQCANCRNKGLPFTCKAFVFDK 60
DB :|||||
DB 1 QKRRNTTHERPKSAKTLIKIDPALKIKTKKVNADQCANCRNKGLPFTCKAFVFDK 60
QY 61 ARKQCLWFPFNSMSSGVKKEFGHEFDLYENKDYIRNCIIIGKRSYKGTVITKSGIKCOP 120
DB :|||||
DB 61 ARKQCLWFPFNSMSSGVKKEFGHEFDLYENKDYIRNCIIIGKRSYKGTVITKSGIKCOP 120
QY 121 WSSMIPHEHSFLPSSYRGKDLQENYCRNPRGEGGPGWCFTSNPEVRYEYCDIPQCSEVEC 180
DB :|||||
DB 121 WSSMIPHEHSFLPSSYRGKDLQENYCRNPRGEGGPGWCFTSNPEVRYEYCDIPQCSEVEC 180
QY 181 MTCNGESYRGLMDHTESGKICQWHDQTPHRRHKFLPERYPDKGFDNCRNPDGQPRPWC 240

```

Db 181 WTCNGESYRGLMDHTEGKICQWDHQTTPHKKFLPERYPDKGDDNYCRNPDGQPRWC 240
QY 241 YLDPHTRWEYCAIKTCADNTMDTDPVLETTETCIQGGEGYRGTVNTIWNIGPCQWRWS 300
Db 241 YLDPHTRWEYCAIKTCADNTMDTDPVLETTETCIQGGEGYRGTVNTIWNIGPCQWRWS 300
QY 301 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSOIPNCDSHGGDC 360
Db 301 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSOIPNCDSHGGDC 360
QY 361 YRGNKKNYMGNSLQTSRGLTCSMDKKNMEDLHRHIFWEPDASKLNENYCRNPDHDAHPW 420
Db 361 YRGNKKNYMGNSLQTSRGLTCSMDKKNMEDLHRHIFWEPDASKLNENYCRNPDHDAHPW 420
QY 421 CYTGNPLIPWDYCPISRCBGDTTPTIV 447
Db 421 CYTGNPLIPWDYCPISRCBGDTTPTIV 447

RESULT 5
US-07-815-333A-2
; Sequence 2, Application US/07815333A
; Patent No. 5342831
; GENERAL INFORMATION:
; APPLICANT: Nakamura, Toshikazu
; APPLICANT: Matsumoto, Kunio
; TITLE OF INVENTION: EPITHELIOCYTE GROWTH ACCELERATOR
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Leydig, Volt & Mayer
; STREET: Two Prudential Plaza, Suite 4900
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: US/07/815,333A
; FILING DATE: 19911227
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Green, Robert F.
; REGISTRATION NUMBER: 27555
; REFERENCE/DOCKET NUMBER: 44069
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 616-5600
; TELEFAX: (312) 616-5700
; TELEX: 25-3533
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 728 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-815-333A-2

Query Match 99.9%; Score 2601; DB 1; Length 728;
Best Local Similarity 99.8%; Pred. No. 1.2e-214;
Matches 446; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 ERKRRNTHEPKSAKTTLLKIDPALKIKTKVNTADQCANRCTRNRKGLPFTCKAFVFDK 60
Db 32 ORKRRNTHEPKSAKTTLLKIDPALKIKTKVNTADQCANRCTRNRKGLPFTCKAFVFDK 91
QY 61 ARKQCLWFFPNSSGSKVKEGHEFDLYENKDYIRNCIIIGKGRSYKGTVTSITSGIKCQP 120
Db 92 ARKQCLWFFPNSSGSKVKEGHEFDLYENKDYIRNCIIIGKGRSYKGTVTSITSGIKCQP 151

QY 121 WSMNIPHEHFLPSSYRGKDLQENYCRNPRGEGGPGWCFTSNPEVRYEYCDIPQCSEVEC 180
Db 152 WSMNIPHEHFLPSSYRGKDLQENYCRNPRGEGGPGWCFTSNPEVRYEYCDIPQCSEVEC 211
QY 181 MTCNGESYRGLMDHTEGKICQWDHQTTPHKKFLPERYPDKGDDNYCRNPDGQPRWC 240
Db 212 MTCNGESYRGLMDHTEGKICQWDHQTTPHKKFLPERYPDKGDDNYCRNPDGQPRWC 271
QY 241 YLDPHTRWEYCAIKTCADNTMDTDPVLETTETCIQGGEGYRGTVNTIWNIGPCQWRWS 300
Db 272 YLDPHTRWEYCAIKTCADNTMDTDPVLETTETCIQGGEGYRGTVNTIWNIGPCQWRWS 331
QY 301 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSOIPNCDSHGGDC 360
Db 332 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSOIPNCDSHGGDC 391
QY 361 YRGNKKNYMGNSLQTSRGLTCSMDKKNMEDLHRHIFWEPDASKLNENYCRNPDHDAHPW 420
Db 392 YRGNKKNYMGNSLQTSRGLTCSMDKKNMEDLHRHIFWEPDASKLNENYCRNPDHDAHPW 451
QY 421 CYTGNPLIPWDYCPISRCBGDTTPTIV 447
Db 452 CYTGNPLIPWDYCPISRCBGDTTPTIV 478

RESULT 6
US-08-087-783A-22
; Sequence 22, Application US/08087783A
; Patent No. 5547856
; GENERAL INFORMATION:
; APPLICANT: Godowski, Paul J., Lokker, Nathalie A., Mark, Melanie R.
; TITLE OF INVENTION: HEPATOCYTE GROWTH FACTOR VARIANTS
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Winpatin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/087,783A
; FILING DATE: 13-Jul-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/884811
; FILING DATE: 18-MAY-92
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/885971
; FILING DATE: 18-MAY-92
; ATTORNEY/AGENT INFORMATION:
; NAME: Marschang, Diane L.
; REGISTRATION NUMBER: 35,600
; REFERENCE/DOCKET NUMBER: P0755779P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-5416
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 728 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
US-08-087-783A-22

Query Match 99.9%; Score 2601; DB 1; Length 728;
Best Local Similarity 99.8%; Pred. No. 1.2e-214;
Matches 446; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 ERKRRNTIHEFKSAKTLTIKIDPALKIKTKKVNADQCANCRTRNKGLPFTCKAFVFDK 60
:|||||
Db 32 QKRRNTIHEFKSAKTLTIKIDPALKIKTKKVNADQCANCRTRNKGLPFTCKAFVFDK 91
:|||||
QY 61 ARKQCLWPFNSMSGVKKFEGHEFDLYENKDYIRNCIIGKGRSYKGTYSITKSGIKCOP 120
:|||||
Db 92 ARKQCLWPFNSMSGVKKFEGHEFDLYENKDYIRNCIIGKGRSYKGTYSITKSGIKCOP 151
:|||||
QY 121 WSSMIPHEHSFLPSSYRGKDLQENCRNPRGEGGPGWCFSTSNPEVRYEYCDIPQCSEVEC 180
:|||||
Db 152 WSSMIPHEHSFLPSSYRGKDLQENCRNPRGEGGPGWCFSTSNPEVRYEYCDIPQCSEVEC 211
:|||||
QY 181 MTCNGESYRGLMDHTESGKICORWDHQTTPHRRKFLPERYPDKGFDNCRNPDGQPRPMC 240
:|||||
Db 212 MTCNGESYRGLMDHTESGKICORWDHQTTPHRRKFLPERYPDKGFDNCRNPDGQPRPMC 271
:|||||
QY 241 YTLDPHTRWEYCAIKTCADNTMNDTDVPLETTECIQGQEGYRGVTNTIWNIGIPCORWDS 300
:|||||
Db 272 YTLDPHTRWEYCAIKTCADNTMNDTDVPLETTECIQGQEGYRGVTNTIWNIGIPCORWDS 331
:|||||
QY 301 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 360
:|||||
Db 332 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 391
:|||||
QY 361 YRGNGKNYMGNLQSOTRSGLTCSMWKDNMEDLHRHIFWEPDASKLNEYCRNPDHDAHGFW 420
:|||||
Db 392 YRGNGKNYMGNLQSOTRSGLTCSMWKDNMEDLHRHIFWEPDASKLNEYCRNPDHDAHGFW 451
:|||||
QY 421 CYTGNPLIPWDYCPISRCBGDTTPTIV 447
:|||||
Db 452 CYTGNPLIPWDYCPISRCBGDTTPTIV 478
:|||||

RESULT 7
US-07-838-410-1
: Sequence 1, Application US/07838410
: Patent No. 5328836
: GENERAL INFORMATION:
: APPLICANT: SHIMA, NOBUYUKI
: APPLICANT: HIGASHIO, KANJI
: APPLICANT: NAGAO, MASAYA
: APPLICANT: OOGAKI, FUMIKO
: APPLICANT: TAKAOKA, HIROAKI
: APPLICANT: TSUDA, ELSUKE
: TITLE OF INVENTION: PLASMIDS CONTAINING DNA ENCODING THE
: TITLE OF INVENTION: AMINO ACID SEQUENCE OF TCF-II, TRANSFORMED CELLS WITH THE
: TITLE OF INVENTION: PLASMIDS, AND PRODUCTION METHOD OF BIOLOGICALLY ACTIVE
: TITLE OF INVENTION: SUBSTANCE USING THE TRANSFORMED CELLS
: NUMBER OF SEQUENCES: 11
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: MILLEN, WHITE, ZELANO & BRANIGAN, P.C.
: STREET: 2200 CLARENDON BLVD.
: CITY: ARLINGTON
: STATE: VA
: COUNTRY: USA
: ZIP: 22201
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/07/838,410
: FILING DATE: 19920311
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: PCT/JP91/00942
: FILING DATE: 15-JUL-1991
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: JP 185852-1990
: FILING DATE: 13-JUL-1990
: ATTORNEY/AGENT INFORMATION:
```

```
: NAME: WHITE, JOHN L.
: REGISTRATION NUMBER: 17,746
: REFERENCE/DOCKET NUMBER: WAK 110
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 703-243-6333
: TELEFAX: 703-243-6410
: TELEX: 64191
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 723 amino acids
: TYPE: AMINO ACID
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: ORIGINAL SOURCE:
: ORGANISM: Homo sapiens
: CELL TYPE: Fibroblast
: FEATURE:
: NAME/KEY: Domain
: LOCATION: 393..405
: OTHER INFORMATION: /note= "INTERNAL AMINO ACID
: OTHER INFORMATION: SEQUENCE IN ALPHA-CHAIN"
: FEATURE:
: NAME/KEY: Protein
: LOCATION: 490..505
: OTHER INFORMATION: /note= "N-TERMINAL AMINO ACID
: OTHER INFORMATION: SEQUENCE OF BETA-CHAIN"
: FEATURE:
: NAME/KEY: Domain
: LOCATION: 605..623
: OTHER INFORMATION: /note= "INTERNAL AMINO ACID
: OTHER INFORMATION: SEQUENCE IN BETA-CHAIN"
: US-07-838-410-1
```

Query Match 98.4%; Score 2563.5; DB 1; Length 723;

Best Local Similarity 98.7%; Pred No. 2e-211; Mismatches 1; Gaps 1;

Matches 441; Conservative 1; Indels 5;

```
QY 1 ERKRRNTIHEFKSAKTLTIKIDPALKIKTKKVNADQCANCRTRNKGLPFTCKAFVFDK 60
:|||||
Db 32 QKRRNTIHEFKSAKTLTIKIDPALKIKTKKVNADQCANCRTRNKGLPFTCKAFVFDK 91
:|||||
QY 61 ARKQCLWPFNSMSGVKKFEGHEFDLYENKDYIRNCIIGKGRSYKGTYSITKSGIKCOP 120
:|||||
Db 92 ARKQCLWPFNSMSGVKKFEGHEFDLYENKDYIRNCIIGKGRSYKGTYSITKSGIKCOP 151
:|||||
QY 121 WSSMIPHEHSFLPSSYRGKDLQENCRNPRGEGGPGWCFSTSNPEVRYEYCDIPQCSEVEC 180
:|||||
Db 152 WSSMIPHEH-----SYRGKDLQENCRNPRGEGGPGWCFSTSNPEVRYEYCDIPQCSEVEC 206
:|||||
QY 181 MTCNGESYRGLMDHTESGKICORWDHQTTPHRRKFLPERYPDKGFDNCRNPDGQPRPMC 240
:|||||
Db 207 MTCNGESYRGLMDHTESGKICORWDHQTTPHRRKFLPERYPDKGFDNCRNPDGQPRPMC 266
:|||||
QY 241 YTLDPHTRWEYCAIKTCADNTMNDTDVPLETTECIQGQEGYRGVTNTIWNIGIPCORWDS 300
:|||||
Db 267 YTLDPHTRWEYCAIKTCADNTMNDTDVPLETTECIQGQEGYRGVTNTIWNIGIPCORWDS 326
:|||||
QY 301 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 360
:|||||
Db 327 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 386
:|||||
QY 361 YRGNGKNYMGNLQSOTRSGLTCSMWKDNMEDLHRHIFWEPDASKLNEYCRNPDHDAHGFW 420
:|||||
Db 387 YRGNGKNYMGNLQSOTRSGLTCSMWKDNMEDLHRHIFWEPDASKLNEYCRNPDHDAHGFW 446
:|||||
QY 421 CYTGNPLIPWDYCPISRCBGDTTPTIV 447
:|||||
Db 447 CYTGNPLIPWDYCPISRCBGDTTPTIV 473
:|||||
```

RESULT 8
US-08-290-937B-1

Query Match	98.4%	Score	2563.5;	DB 1;	Length	723;			
Best Local Similarity	98.7%;	Pred. No.	2e-211;						
Matches	441;	Conservative	1;	Mismatches	0;	Indels	5;	Gaps	1;
Qy	1	ERKRNTTHEFKKSAKTTLIKIDPALKIKTKYNTADOCANCRTRNKGLPFTCKAFVDDK	60						
Db	32	QKRKRNTTHEFKKSAKTTLIKIDPALKIKTKYNTADOCANCRTRNKGLPFTCKAFVDDK	91						
Qy	61	ARKQCLWFPFNSMSGVKKEGFHFDLYENKDYIRNCIIGKGRSYKGTVTSITKSGIKQCP	120						
Db	92	ARKQCLWFPFNSMSGVKKEGFHFDLYENKDYIRNCIIGKGRSYKGTVTSITKSGIKQCP	151						
Qy	121	WSMTPHEHSFLPSSYRGKDIQENYCRNPGEEGGPMCFTSNPEVRYEVCDDIQOSEVBC	180						
Db	152	WSMTPHEH-----SYRGKDIQENYCRNPGEEGGPMCFTSNPEVRYEVCDDIQOSEVBC	206						
Qy	181	MTNCGESYGLMDHTESGKICQRWDHQTPIRRHKFLPERYPDKGFDDBNYCRNPDGQPRWC	240						
Db	207	MTNCGESYGLMDHTESGKICQRWDHQTPIRRHKFLPERYPDKGFDDBNYCRNPDGQPRWC	266						
Qy	241	YTLDPHTRWEYCAIKTCADNTMTNDTVPLETTCIOGGEGYRGTVNTTWNGIPCORWDS	300						
Db	267	YTLDPHTRWEYCAIKTCADNTMTNDTVPLETTCIOGGEGYRGTVNTTWNGIPCORWDS	326						
Qy	301	QYPHEHDMTPENFKCKDLRENYCRNPDSGSPWCFITDPNIRVGYCSQIPNCNDSHGDCD	360						

Query Match	98.4%	Score	2563.5	DB 1	Length	723			
Best Local Similarity	98.7%	Pred. No.	2e-211						
Matches	441	Conservative	1	Mismatches	0	Indels	5	Gaps	1
Qy	1	ERKRNTIHEPKKSAKTTLIKIDPALKTKTKVNTADQCANCRCTRNGKLPETCKAFVFDK	60						
		:							
Db	32	QKRKNTIHEEKKSAKTTLIKIDPALKTKTKVNTADQCANCRCTRNGKLPETCKAFVFDK	91						
Qy	61	ARKOCLWPPFNMSGSKYKKEGFHEFDLYENKDYIRNCLIIIGKRSYKGTVTSTKSGIKCQP	120						
		:							
Db	92	ARKOCLWPPFNMSGSKYKKEGFHEFDLYENKDYIRNCLIIIGKRSYKGTVTSTKSGIKCQP	151						
Qy	121	WSSMIPBHSFLPSYRGKDLOENYCRNPRGEGGPNCFSTNDEVRYTEVCDIPOCSEVEC	180						

Db 152 WSSMIPHEH-----SYRGKDLQENYCRNPRGEGGPGWCFSTSNPEVRYEYCDIPQCSEVEC 206
QY 181 MTCNGESYRGLMDHTESGKICORWDHQTTPHRRKFLPERYPDKGFDNCRNPDGQPRPWC 240
Db 207 MTCNGESYRGLMDHTESGKICORWDHQTTPHRRKFLPERYPDKGFDNCRNPDGQPRPWC 266
QY 241 YTLDPHTRWYCAITCADNTMNDTVDPLETTECIGQGEGYRGVTNTIWNIGIPQCRWDS 300
Db 267 YTLDPHTRWYCAITCADNTMNDTVDPLETTECIGQGEGYRGVTNTIWNIGIPQCRWDS 326
QY 301 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 360
Db 327 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 386
QY 361 YRGNGKNTMGNLSQTRSGLTCSMWKDNMEDLHRHIFWEPDASKLNYCRNPDGQPRPWC 420
Db 387 YRGNGKNTMGNLSQTRSGLTCSMWKDNMEDLHRHIFWEPDASKLNYCRNPDGQPRPWC 446
QY 421 CYTGNPLIPWDYCPISRCBGDTTPTIV 447
Db 447 CYTGNPLIPWDYCPISRCBGDTTPTIV 473

RESULT 10

US-08-404-643-1

; Sequence 1, Application US/08404643

; Patent No. 5658742

; GENERAL INFORMATION:

; APPLICANT: HIGASHIO, KANJI

; APPLICANT: SHIMA, NOBUYUKI

; APPLICANT: OOGAKI, FUMIKO

; TITLE OF INVENTION: MONOCLONAL ANTIBODY

; NUMBER OF SEQUENCES: 1

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: PATENT ADMINISTRATOR, TESTA, HURWITZ &

; ADDRESSEE: THIBEAULT

; STREET: 53 STATE STREET

; CITY: BOSTON

; STATE: MA

; COUNTRY: USA

; ZIP: 02109

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/404,643

; FILING DATE: 15-MAR-1995

; CLASSIFICATION: 436

; ATTORNEY/AGENT INFORMATION:

; NAME: CAMPBELL, PAULA A

; REGISTRATION NUMBER: 32,503

; REFERENCE/DOCKET NUMBER: FJN-034 (3999/35)

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (617)-248-7000

; TELEFAX: (617)-248-7100

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 723 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-08-404-643-1

.Query Match

Best Local Similarity 98.4%; Score 2563.5; DB 1; Length 723;

Matches 441; Conservative 1; Mismatches 0; Indels 5; Gaps 1;

QY 1

ERKRRNTHEFKKSAKTLTIKIDPALKIKTKKVNADQCANRCKTRNKGLPFTCKAFVFDK 60

:|||||

Db 32 QKRRNTHEFKKSAKTLTIKIDPALKIKTKKVNADQCANRCKTRNKGLPFTCKAFVFDK 91
QY 61 ARKQCLWFPFNSMSGVKKEFGHEFDLYENKDYIRNCIIIGKGRSYKGTYSITKSGIKCOP 120
Db 92 ARKQCLWFPFNSMSGVKKEFGHEFDLYENKDYIRNCIIIGKGRSYKGTYSITKSGIKCOP 151
QY 121 WSSMIPHEHFLPSSYRGKDLQENYCRNPRGEGGPGWCFSTSNPEVRYEYCDIPQCSEVEC 180
Db 152 WSSMIPHEH-----SYRGKDLQENYCRNPRGEGGPGWCFSTSNPEVRYEYCDIPQCSEVEC 206
QY 181 MTCNGESYRGLMDHTESGKICORWDHQTTPHRRKFLPERYPDKGFDNCRNPDGQPRPWC 240
Db 207 MTCNGESYRGLMDHTESGKICORWDHQTTPHRRKFLPERYPDKGFDNCRNPDGQPRPWC 266
QY 241 YTLDPHTRWYCAITCADNTMNDTVDPLETTECIGQGEGYRGVTNTIWNIGIPQCRWDS 300
Db 267 YTLDPHTRWYCAITCADNTMNDTVDPLETTECIGQGEGYRGVTNTIWNIGIPQCRWDS 326
QY 301 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 360
Db 327 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 386
QY 361 YRGNGKNTMGNLSQTRSGLTCSMWKDNMEDLHRHIFWEPDASKLNYCRNPDGQPRPWC 420
Db 387 YRGNGKNTMGNLSQTRSGLTCSMWKDNMEDLHRHIFWEPDASKLNYCRNPDGQPRPWC 446
QY 421 CYTGNPLIPWDYCPISRCBGDTTPTIV 447
Db 447 CYTGNPLIPWDYCPISRCBGDTTPTIV 473

RESULT 11

US-09-194-326-1

; Sequence 1, Application US/09194326

; Patent No. 6306827

; GENERAL INFORMATION:

; APPLICANT: Kinoshita, Masahiko

; APPLICANT: Ogawa, Hiroshi

; APPLICANT: Masanaga, Hiroaki

; APPLICANT: Kobayashi, Fumie

; APPLICANT: Yamaguchi, Kyoji

; APPLICANT: Higashio, Kanji

; TITLE OF INVENTION: Agent for Preventing and/or Treating Renal Disease

; FILE REFERENCE: FJN-069

; CURRENT APPLICATION NUMBER: US/09/194,326

; EARLIER FILING DATE: 1998-11-24

; EARLIER APPLICATION NUMBER: PCT/JP98/01221

; EARLIER FILING DATE: 1998-03-20

; EARLIER APPLICATION NUMBER: JP 94989

; EARLIER FILING DATE: 1997-03-28

; NUMBER OF SEQ ID NOS: 3

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 1

; LENGTH: 723

; TYPE: PRT

; ORGANISM: Homo sapiens

; FEATURE:

; OTHER INFORMATION: TCF-II

US-09-194-326-1

Query Match

Best Local Similarity 98.4%; Score 2563.5; DB 4; Length 723;

Matches 441; Conservative 1; Mismatches 0; Indels 5; Gaps 1;

QY 1

ERKRRNTHEFKKSAKTLTIKIDPALKIKTKKVNADQCANRCKTRNKGLPFTCKAFVFDK 60

:|||||

Db 32 QKRRNTHEFKKSAKTLTIKIDPALKIKTKKVNADQCANRCKTRNKGLPFTCKAFVFDK 91

QY 61 ARKQCLWFPFNSMSGVKKEFGHEFDLYENKDYIRNCIIIGKGRSYKGTYSITKSGIKCOP 120

Db 92 ARKQCLWFPFNSMSGVKKEFGHEFDLYENKDYIRNCIIIGKGRSYKGTYSITKSGIKCOP 151

QY 121 WSSMIPHEHFLPSSYRGKDLQENYCRNPRGEGGPGWCFSTSNPEVRYEYCDIPQCSEVEC 180

		Matches. 440:	Conservative	2:	Mismatches	0:	Indels	5:	Gaps	1:
Db	152	WSSMIPHEH-----SYRGKDLQENYCRNPRGEGPWCFTSNPEVRYEVCDIPOCSEVEC	206							
Qy	181	MTCNGESYRGLMDHTESGKICQWRDQHPHRRKFLPERYPDKGFDNDYCRNPDGQPRPWC	240							
Db	207	MTCNGESYRGLMDHTESGKICQWRDQHPHRRKFLPERYPDKGFDNDYCRNPDGQPRPWC	266							
Qy	241	YTLDPHTRWEYCAIKTCADNTMNDTVPLETTTCIOGEGYRGTVNTIWNIGIPCORWDS	300							
Db	267	YTLDPHTRWEYCAIKTCADNTMNDTVPLETTTCIOGEGYRGTVNTIWNIGIPCORWDS	326							
Qy	301	QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC	360							
Db	327	QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC	386							
Qy	361	YRGNKGYMGNLSQTRSGLTCSMWDKNNMEDLHRHIFWEPDASKLNNYCRNPDHDAHPW	420							
Db	387	YRGNKGYMGNLSQTRSGLTCSMWDKNNMEDLHRHIFWEPDASKLNNYCRNPDHDAHPW	446							
Qy	421	CYTGNPLIPWDYCPISRCGEGDTTPTIV	447							
Db	447	CYTGNPLIPWDYCPISRCGEGDTTPTIV	473							

```
RESULT 13
US-08-290-937B-3
; Sequence 3, Application US/08290937B
; Patent No. 5648233
; GENERAL INFORMATION:
; APPLICANT: YAMAGUCHI, KYOJI
; APPLICANT: SHIMA, NOBUYUKI
; APPLICANT: MURAKAMI, AKIHIKO
; APPLICANT: GOTO, MASAOKI
; APPLICANT: TSUDA, EISUKE
; APPLICANT: MASUNAGA, HIROAKI
; APPLICANT: TAKAHARA, REIKO
; APPLICANT: OOGAKI, FUMIKO
; APPLICANT: UEDA, MASATSUGU
; APPLICANT: HIGASHIO, KANJI
; TITLE OF INVENTION: MODIFIED TCF
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Testa, Hurwitz & Thibault
; STREET: 125 High St.
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/290,937B
; FILING DATE: 19-AUG-1994
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: CAMPBELL, PAULA A.
; REGISTRATION NUMBER: 32,503
; REFERENCE/DOCKET NUMBER: FJN-022
; TELECOMMUNICATION INFORMATION:
```

```

; TELEPHONE: (617) 248-7000
; TELEFAX: (617) 248-7100
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 723 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-290-937B-3

Query Match      98.2%; Score 2557.5; DB 1; Length 723;
Best Local Similarity 98.4%; Pred. No. 6.5e-211;
Matches 440; Conservative 1; Mismatches 1; Indels 5; Gaps 1;

QY 1 ERKRNTIHEFKSAKTTLIKIDPALKIKTKKVTADQCANRCTRNRKGLPFTCKAFVFDK 60
DB 32 QRKRNTIHEFKSAKTTLIKIDPALKIKTKKVTADQCANRCTRNRKGLPFTCKAFVFDK 91
QY 61 ARKQCLWFPNMSGGVKKEGHEFDLYENKDYIRNCIIIGKRSYKGTVSTTKSGIKQCP 120
DB 92 ARKQCLWFPNMSGGVKKEGHEFDLYENKDYIRNCIIIGKRSYKGTVSTTKSGIKQCP 151
QY 121 WSSMIPHEHFLPSSYRGKDLQENYCRNPRGEGGWCFTSNPEVRYEVCIDIPOCSEVEC 180
DB 152 WSSMIPHEH-----SYRGKDLQENYCRNPRGEGGWCFTSNPEVRYEVCIDIPOCSEVEC 206
QY 181 MTCNGESYRGLMDHTESGKICQWDHQTPHRHKFLPERYPDKGFDNNYCRNPDGQPRPWC 240
DB 207 MTCNGESYRGLMDHTESGKICQWDHQTPHRHKFLPERYPDKGFDNNYCRNPDGQPRPWC 266
QY 241 YLDPHTRWEYCAIKTCADNTMNDTVPLETTETCIGQGGEGYRGTVNTIWNIGIPCQRWDS 300
DB 267 YLDPHTRWEYCAIKTCADNTMNDTVPLETTETCIGQGGEGYRGTVNTIWNIGIPCQRWDS 326
QY 301 QYPHEHMTPEHFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 360
DB 327 QYPHEHMTPEHFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 386
QY 361 YRGNKKNYMGNSQTRSGLTCSMWDKNMEDLHRHIFWEPDASKLNENYCRNPDHDAHGCPW 420
DB 387 YRGNKKNYMGNSQTRSGLTCSMWDKNMEDLHRHIFWEPDASKLNENYCRNPDHDAHGCPW 446
QY 421 CYTGNPLIPWDYCPISRCGDTTPTIV 447
DB 447 CYTGNPLIPWDYCPISRCGDTTPTIV 473

RESULT 14
US-09-194-326-2
; Sequence 2, Application US/09194326
; Patent No. 6306827
; GENERAL INFORMATION:
; APPLICANT: Kinoshaki, Masahiko
; APPLICANT: Ogawa, Hiromi
; APPLICANT: Masanaga, Hiroaki
; APPLICANT: Kobayashi, Fumie
; APPLICANT: Yamaguchi, Kyoji
; APPLICANT: Higashio, Kanji
; TITLE OF INVENTION: Agent for Preventing and/or Treating Renal Disease
; FILE REFERENCE: FJN-069
; CURRENT APPLICATION NUMBER: US/09/194,326
; CURRENT FILING DATE: 1998-11-24
; EARLIER APPLICATION NUMBER: PCT/JF98/01221
; EARLIER FILING DATE: 1998-03-20
; EARLIER APPLICATION NUMBER: JP 94989
; EARLIER FILING DATE: 1997-03-28
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn ver. 2.0
; SEQ ID NO 2
; LENGTH: 723
; TYPE: PRT
; ORGANISM: Artificial Sequence

```

```

; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: RKRR2AAAA
; OTHER INFORMATION: mutant of TCF-II
US-09-194-326-2

Query Match      97.6%; Score 2541.5; DB 4; Length 723;
Best Local Similarity 98.9%; Pred. No. 1.5e-209;
Matches 437; Conservative 0; Mismatches 0; Indels 5; Gaps 1;

QY 6 NTIHEFKSAKTTLIKIDPALKIKTKKVTADQCANRCTRNRKGLPFTCKAFVFDKARKQC 65
DB 37 NTIHEFKSAKTTLIKIDPALKIKTKKVTADQCANRCTRNRKGLPFTCKAFVFDKARKQC 96
QY 66 LWFPNMSGGVKKEGHEFDLYENKDYIRNCIIIGKRSYKGTVSTTKSGIKQCPWSSMI 125
DB 97 LWFPNMSGGVKKEGHEFDLYENKDYIRNCIIIGKRSYKGTVSTTKSGIKQCPWSSMI 156
QY 126 PHEHSLFPLSSYRGKDLQENYCRNPRGEGGWCFTSNPEVRYEVCIDIPOCSEVECMTGNG 185
DB 157 PHEH-----SYRGKDLQENYCRNPRGEGGWCFTSNPEVRYEVCIDIPOCSEVECMTGNG 211
QY 186 ESYRGLMDHTESGKICQWDHQTPHRHKFLPERYPDKGFDNNYCRNPDGQPRPWCYTLDLP 245
DB 212 ESYRGLMDHTESGKICQWDHQTPHRHKFLPERYPDKGFDNNYCRNPDGQPRPWCYTLDLP 271
QY 246 HTRWEYCAIKTCADNTMNDTVPLETTETCIGQGGEGYRGTVNTIWNIGIPCQRWDSQYPHE 305
DB 272 HTRWEYCAIKTCADNTMNDTVPLETTETCIGQGGEGYRGTVNTIWNIGIPCQRWDSQYPHE 331
QY 306 HDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDCYRNG 365
DB 332 HDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDCYRNG 391
QY 366 KNYMGNSQTRSGLTCSMWDKNMEDLHRHIFWEPDASKLNENYCRNPDHDAHGCPWCYTGN 425
DB 392 KNYMGNSQTRSGLTCSMWDKNMEDLHRHIFWEPDASKLNENYCRNPDHDAHGCPWCYTGN 451
QY 426 PLIPWDYCPISRCGDTTPTIV 447
DB 452 PLIPWDYCPISRCGDTTPTIV 473

RESULT 15
US-08-700-519J-18
; Sequence 18, Application US/08700519J
; Patent No. 6399744
; GENERAL INFORMATION:
; APPLICANT: Kinoshaki, Masahiko
; APPLICANT: Yamaguchi, Kyoji
; APPLICANT: Goto, Masaaki
; APPLICANT: Murakami, Akihiko
; APPLICANT: Ueda, Masatsugu
; APPLICANT: Higashio, Kanji
; APPLICANT: Yamashita, Yasushi
; TITLE OF INVENTION: TCF MUTANT
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESS: Burgess, Ryan and Wayne
; STREET: 370 Lexington Avenue, Suite 2105
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 10017
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3 1/5 inch diskette
; COMPUTER: PC'S LIMITED SYSTEM 200
; OPERATING SYSTEM: WORD FOR WINDOWS
; SOFTWARE: MICROSOFT WINDOWS 98
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/700,519J
; FILING DATE: 26-Aug-1996
; CLASSIFICATION: <unknown>
; PRIOR APPLICATION DATA:

```

```
APPLICATION NUMBER: <Unknown>
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Wayne, Milton J.
REGISTRATION NUMBER: 17,906
REFERENCE/DOCKET NUMBER: U-Wp-5240PCT
TELEPHONE: 212-683-8150
TELEFAX: 212-532-4285
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 723
TYPE: AMINO ACID
STRANDEDNESS: SINGLE
TOPOLOGY: LINEAR
SEQUENCE DESCRIPTION: SEQ ID NO: 18:
US-08-700-519J-18

Query Match          97.6%; Score 2541.5; DB 4; Length 723;
Best Local Similarity 98.9%; Pred. No. 1.5e+209;
-Matches 437; Conservative 0; Mismatches 0; Indels 5; Gaps 1;

QY 6 NTIHEFKSAKTLIKIDPALKIKTKKYNTADQCANCRTRNKGLPFTCKAFVFDKARKQC 65
DB 37 NTIHEFKSAKTLIKIDPALKIKTKKYNTADQCANCRTRNKGLPFTCKAFVFDKARKQC 96
QY 66 LWFPNSMSGYKKEFGHEFDLYENKDYIRNCIIIGKGRSYKGTYSITKSGIKCOPWSSMI 125
DB 97 LWFPNSMSGYKKEFGHEFDLYENKDYIRNCIIIGKGRSYKGTYSITKSGIKCOPWSSMI 156
QY 126 PHEHFLPSSYRGKDLQENYCRNPRGEGGPMWCFSTNPEVRYEVCIDIQCSEVECMTCNG 185
DB 157 PHEH-----SYRGKDLQENYCRNPRGEGGPMWCFSTNPEVRYEVCIDIQCSEVECMTCNG 211
QY 186 ESYRGLMDHTESGKICQWDHQTPIRHHKFLPERYPDKGFDNNYCRNPDGQRPWCYTLDLP 245
DB 212 ESYRGLMDHTESGKICQWDHQTPIRHHKFLPERYPDKGFDNNYCRNPDGQRPWCYTLDLP 271
QY 246 HTRWEYCAIKTCADNTMNDTDVPLETTECIGGEGYRGTVNTIWNIGPCQWDSQYPHE 305
DB 272 HTRWEYCAIKTCADNTMNDTDVPLETTECIGGEGYRGTVNTIWNIGPCQWDSQYPHE 331
QY 306 HDMTPENFKCKDLRENYCRNPDGSESPWCFSTDPTNIRVGYCSQIPNCDMSHGQDCYRGNG 365
DB 332 HDMTPENFKCKDLRENYCRNPDGSESPWCFSTDPTNIRVGYCSQIPNCDMSHGQDCYRGNG 391
QY 366 KNTWGNLSQTRSLGTCMWDKNNEDLHRHIFWEPDASKLNNYCRNPDHAGHPWCYTGN 425
DB 392 KNTWGNLSQTRSLGTCMWDKNNEDLHRHIFWEPDASKLNNYCRNPDHAGHPWCYTGN 451
QY 426 PLIPWDYCPISRCGDTTPTIV 447
DB 452 PLIPWDYCPISRCGDTTPTIV 473
```

Search completed: April 25, 2003, 15:45:23
Job time : 23 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 25, 2003, 15:44:38 ; Search time 21 seconds
(without alignments)
1705.620 Million cell updates/sec

Title: US-09-674-377B-1

Perfect score: 2604

Sequence: 1 ERKRNTIHEFKSAKTTLI.....IPWDYCPISRCEGDTTPTIV 447

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 301932 seqs, 80129803 residues

Total number of hits satisfying chosen parameters: 301932

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications_AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2601	99.9	447	9	US-10-081-309-2
2	2601	99.9	728	1	US-08-605-221-2
3	2563.5	98.4	723	1	US-08-605-221-4
4	2541.5	97.6	723	9	US-10-133-912-18
5	2539.5	97.5	723	9	US-10-133-912-19
6	1180.5	45.3	711	9	US-10-133-036-2
7	1059	40.7	812	9	US-09-335-325-1
8	1059	40.7	812	9	US-10-131-241-1
9	1059	40.7	812	10	US-09-788-142-1
10	1059	40.7	812	10	US-09-761-120-1
11	1059	40.7	812	10	US-09-873-676-81
12	1042	40.0	569	10	US-09-946-893-5
13	1042	40.0	571	10	US-09-946-893-8
14	1042	40.0	576	10	US-09-946-893-6
15	1042	40.0	791	9	US-09-967-386-1
16	1042	40.0	810	10	US-09-946-893-2
17	1034	39.7	451	9	US-10-157-369-2
18	1022.5	39.3	378	9	US-09-335-325-41
19	1022.5	39.3	378	9	US-10-131-241-41

20	1022.5	39.3	378	10	US-09-761-120-41	Sequence 41, Appl
21	1003.5	38.5	459	10	US-09-761-120-46	Sequence 46, Appl
22	1003	38.5	352	9	US-09-335-325-39	Sequence 39, Appl
23	1003	38.5	352	9	US-10-131-241-39	Sequence 39, Appl
24	1003	38.5	352	10	US-09-761-120-39	Sequence 39, Appl
25	988.5	38.0	378	9	US-09-335-325-42	Sequence 42, Appl
26	988.5	38.0	378	9	US-10-131-241-42	Sequence 42, Appl
27	988.5	38.0	378	10	US-09-873-676-1	Sequence 1, Appl
28	985	37.8	364	9	US-10-157-369-4	Sequence 4, Appl
29	966.5	37.1	458	10	US-09-946-893-4	Sequence 4, Appl
30	962.5	37.0	368	10	US-09-761-120-42	Sequence 42, Appl
31	960	36.9	352	9	US-09-335-325-40	Sequence 40, Appl
32	960	36.9	352	9	US-10-131-241-40	Sequence 40, Appl
33	960	36.9	352	10	US-09-761-120-40	Sequence 40, Appl
34	942	36.2	339	9	US-09-335-325-2	Sequence 2, Appl
35	942	36.2	339	9	US-10-131-241-2	Sequence 2, Appl
36	942	36.2	339	10	US-09-788-142-2	Sequence 2, Appl
37	942	36.2	339	10	US-09-761-120-2	Sequence 2, Appl
38	940	36.1	339	9	US-09-335-325-5	Sequence 5, Appl
39	940	36.1	339	9	US-10-131-241-5	Sequence 5, Appl
40	940	36.1	339	10	US-09-788-142-5	Sequence 5, Appl
41	940	36.1	339	10	US-09-761-120-5	Sequence 5, Appl
42	929	35.7	423	12	US-10-036-869-23	Sequence 23, Appl
43	922	35.4	339	9	US-09-335-325-6	Sequence 6, Appl
44	922	35.4	339	9	US-10-131-241-6	Sequence 6, Appl
45	922	35.4	339	10	US-09-788-142-6	Sequence 6, Appl

ALIGNMENTS

RESULT 1

US-10-081-309-2

; Sequence 2, Application US/10081309

; Publication No. US20030012775A1

; GENERAL INFORMATION:

; APPLICANT: Hoffmann-La Roche Inc.

; TITLE OF INVENTION: PEG Conjugates of NK4

; FILE REFERENCE: 20859

; CURRENT APPLICATION NUMBER: US/10/081.309

; CURRENT FILING DATE: 2002-06-04

; NUMBER OF SEQ ID NOS: 2

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 2

; LENGTH: 447

; TYPE: PRT

; ORGANISM: homosapiens

US-10-081-309-2

Query Match 99.9%; Score 2601; , DB 9; Length 447;

Best Local Similarity 99.8%; Pred. No. 6.4e-202;

Matches 446; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy	1	ERKRNTIHEFKSAKTTLIKIDPALKIKTKKVNADQCANCRNKGLPFTCKAFVDEK	60
Db	1	QRKRNTIHEFKSAKTTLIKIDPALKIKTKKVNADQCANCRNKGLPFTCKAFVDEK	60
Qy	61	ARKQCLWFPFNSMSSGVKKEFGHEFDLYENKDYIRNCIIIGKGRSVKGTVSTIKSGIKCOP	120
Db	61	ARKQCLWFPFNSMSSGVKKEFGHEFDLYENKDYIRNCIIIGKGRSVKGTVSTIKSGIKCOP	120
Qy	121	WSSMIPHEHSFLPSSYRGKDLQENYCRNPRGEGGWCFWCFSTNPEVRYEVCIDIPOCSEVEC	180
Db	121	WSSMIPHEHSFLPSSYRGKDLQENYCRNPRGEGGWCFWCFSTNPEVRYEVCIDIPOCSEVEC	180
Qy	181	MTNCGESYRGLMDHTESGKICORWDHQTPHRRKFLPERYPDKGFDNNYCRNPDGQPRWC	240
Db	181	MTNCGESYRGLMDHTESGKICORWDHQTPHRRKFLPERYPDKGFDNNYCRNPDGQPRWC	240
Qy	241	YTLDPHTRWEYCAIKTCADNTMNDTVDVPLETTECLOGGEGYRGTVNTIWNIPCORWDS	300
Db	241	YTLDPHTRWEYCAIKTCADNTMNDTVDVPLETTECLOGGEGYRGTVNTIWNIPCORWDS	300

```
QY 301 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 360
Db 301 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 360
QY 361 YRGNGKNYMGNSQTRSGLTCSMDKKNMEDLHRHIFWEPDASKLNNENYCRNPDDDAHGWP 420
Db 361 YRGNGKNYMGNSQTRSGLTCSMDKKNMEDLHRHIFWEPDASKLNNENYCRNPDDDAHGWP 420
QY 421 CYTGNPLIPWDYCPISRCGEGDTTPTIV 447
Db 421 CYTGNPLIPWDYCPISRCGEGDTTPTIV 447

RESULT 2
US-08-605-221-2
; Sequence 2, Application US/08605221
; Publication No. US20030060403A1
; GENERAL INFORMATION:
; APPLICANT: NAKAMURA, Toshikazu
; TITLE OF INVENTION: THERAPEUTIC AGENT FOR DISORDER IN BRAIN AND NERVE
; FILE REFERENCE: 2520-0101P
; CURRENT APPLICATION NUMBER: US/08/605,221
; CURRENT FILING DATE: 1996-03-15
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 728
; TYPE: PRT
; ORGANISM: Homo sapiens
; OTHER INFORMATION: HGF-A (Full length HGF)
US-08-605-221-2

Query Match 99.9%; Score 2601; DB 1; Length 728;
Best Local Similarity 99.8%; Pred. No. 1.2e-201;
Matches 446; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ERKRRNTIHEFKKSAKTLTIKIDPALKIKTKKVTADQCANRCTRNGKGLPFTCKAFVFDK 60
Db 32 QKRRNTIHEFKKSAKTLTIKIDPALKIKTKKVTADQCANRCTRNGKGLPFTCKAFVFDK 91
QY 61 ARKQCLWFPFNSMSSGVYKKEFGHEFDLYENKDYIRNCIIIGKGRSYKGTVSTKSGIKQCP 120
Db 92 ARKQCLWFPFNSMSSGVYKKEFGHEFDLYENKDYIRNCIIIGKGRSYKGTVSTKSGIKQCP 151
QY 121 WSSMIPHEHSLPSSYRGKDLQENYCRNPRGEGGWCFTSNPEVRYEVCDDIPQCSEVEC 180
Db 152 WSSMIPHEHSLPSSYRGKDLQENYCRNPRGEGGWCFTSNPEVRYEVCDDIPQCSEVEC 206
QY 181 MTCNGESYRGLMDHTESGKICQRDWDHQPTRHKFLPERYPDKGFDNNYCRNPDGQPRPWC 240
Db 207 MTCNGESYRGLMDHTESGKICQRDWDHQPTRHKFLPERYPDKGFDNNYCRNPDGQPRPWC 266
QY 241 YLDPHTRWECYCAIKTCADNTMNDTVPLETTTCIOGEGYRGTVNTIWNIGIPCORWDS 300
Db 267 YLDPHTRWECYCAIKTCADNTMNDTVPLETTTCIOGEGYRGTVNTIWNIGIPCORWDS 326
QY 301 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 360
Db 327 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 386
QY 361 YRGNGKNYMGNSQTRSGLTCSMDKKNMEDLHRHIFWEPDASKLNNENYCRNPDDDAHGWP 420
Db 387 YRGNGKNYMGNSQTRSGLTCSMDKKNMEDLHRHIFWEPDASKLNNENYCRNPDDDAHGWP 446
QY 421 CYTGNPLIPWDYCPISRCGEGDTTPTIV 447
Db 447 CYTGNPLIPWDYCPISRCGEGDTTPTIV 473

RESULT 3
US-08-605-221-4
; Sequence 4, Application US/08605221
; Publication No. US20030060403A1
```

```
; GENERAL INFORMATION:
; APPLICANT: NAKAMURA, Toshikazu
; TITLE OF INVENTION: THERAPEUTIC AGENT FOR DISORDER IN BRAIN AND NERVE
; FILE REFERENCE: 2520-0101P
; CURRENT APPLICATION NUMBER: US/08/605,221
; CURRENT FILING DATE: 1996-03-15
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 723
; TYPE: PRT
; ORGANISM: Homo sapiens
; OTHER INFORMATION: HGF-B (five amino acids deletion type)
US-08-605-221-4

Query Match 98.4%; Score 2563.5; DB 1; Length 723;
Best Local Similarity 98.7%; Pred. No. 1.2e-198;
Matches 441; Conservative 1; Mismatches 0; Indels 5; Gaps 1;

QY 1 ERKRRNTIHEFKKSAKTLTIKIDPALKIKTKKVTADQCANRCTRNGKGLPFTCKAFVFDK 60
Db 32 QKRRNTIHEFKKSAKTLTIKIDPALKIKTKKVTADQCANRCTRNGKGLPFTCKAFVFDK 91
QY 61 ARKQCLWFPFNSMSSGVYKKEFGHEFDLYENKDYIRNCIIIGKGRSYKGTVSTKSGIKQCP 120
Db 92 ARKQCLWFPFNSMSSGVYKKEFGHEFDLYENKDYIRNCIIIGKGRSYKGTVSTKSGIKQCP 151
QY 121 WSSMIPHEHSLPSSYRGKDLQENYCRNPRGEGGWCFTSNPEVRYEVCDDIPQCSEVEC 180
Db 152 WSSMIPHEHSLPSSYRGKDLQENYCRNPRGEGGWCFTSNPEVRYEVCDDIPQCSEVEC 206
QY 181 MTCNGESYRGLMDHTESGKICQRDWDHQPTRHKFLPERYPDKGFDNNYCRNPDGQPRPWC 240
Db 207 MTCNGESYRGLMDHTESGKICQRDWDHQPTRHKFLPERYPDKGFDNNYCRNPDGQPRPWC 266
QY 241 YLDPHTRWECYCAIKTCADNTMNDTVPLETTTCIOGEGYRGTVNTIWNIGIPCORWDS 300
Db 267 YLDPHTRWECYCAIKTCADNTMNDTVPLETTTCIOGEGYRGTVNTIWNIGIPCORWDS 326
QY 301 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 360
Db 327 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 386
QY 361 YRGNGKNYMGNSQTRSGLTCSMDKKNMEDLHRHIFWEPDASKLNNENYCRNPDDDAHGWP 420
Db 387 YRGNGKNYMGNSQTRSGLTCSMDKKNMEDLHRHIFWEPDASKLNNENYCRNPDDDAHGWP 446
QY 421 CYTGNPLIPWDYCPISRCGEGDTTPTIV 447
Db 447 CYTGNPLIPWDYCPISRCGEGDTTPTIV 473

RESULT 4
US-10-133-912-18
; Sequence 18, Application US/10133912
; Patent No. US20020165358A1
; GENERAL INFORMATION:
; APPLICANT: Kinoshita, Masahiko
; Yamaguchi, Kyoji
; Goto, Masaaki
; Murakami, Akihiko
; Ueda, Masatsugu
; Higashio, Kanji
; Yamashita, Yasushi
; TITLE OF INVENTION: TCF MUTANT
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Burgess, Ryan and Wayne
; STREET: 370 Lexington Avenue, Suite 2105
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 10017
```

COMPUTER READABLE FORM:
MEDIUM TYPE: 3 1/5 inch diskette
COMPUTER: PC'S LIMITED SYSTEM 200
OPERATING SYSTEM: WORD FOR WINDOWS
SOFTWARE: MICROSOFT WINDOWS 98
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/133,912
FILING DATE: 25-Apr-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/700,519
FILING DATE: 26-Aug-1996
ATTORNEY/AGENT INFORMATION:
NAME: Wayne, Milton J.
REFERENCE/DOCKET NUMBER: U-Wp-5240PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-683-8150
TELEFAX: 212-532-4285
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 723
TYPE: AMINO ACID
STRANDEDNESS: SINGLE
TOPOLOGY: LINEAR
SEQUENCE DESCRIPTION: SEQ ID NO: 18:
US-10-133-912-18

Query Match 97.6%; Score 2541.5; DB 9; Length 723;
Best Local Similarity 98.9%; Pred. No. 7.2e-197;
Matches 437; Conservative 0; Mismatches 0; Indels 5; Gaps 1;
QY 6 NTIHEFKSAKTLTIKIDPALKIKTKVNTADQCANRCKTRNKGLPFTCKAFVDFKARQC 65
Db 37 NTIHEFKSAKTLTIKIDPALKIKTKVNTADQCANRCKTRNKGLPFTCKAFVDFKARQC 96
QY 66 LWTFPNSMSGVKEFGHEFDLYENKDYIRNCIIIGKRSYKGTVSITKSGIKQCPWSSMI 125
Db 97 LWTFPNSMSGVKEFGHEFDLYENKDYIRNCIIIGKRSYKGTVSITKSGIKQCPWSSMI 156
QY 126 PHHSFLPSSYRGKDLQENYCRNPRGEGGPGWCFSTSNPEVRYEVCIDIPOCSEVECMTCNG 185
Db 157 PHEH-----SYRGKDLQENYCRNPRGEGGPGWCFSTSNPEVRYEVCIDIPOCSEVECMTCNG 211
QY 186 ESYRGLMDHTESGKICQWHDQTPHRRKFLPERYPDKGFDNDYCRNPDGQPRPWCYTLDP 245
Db 212 ESYRGLMDHTESGKICQWHDQTPHRRKFLPERYPDKGFDNDYCRNPDGQPRPWCYTLDP 271
QY 246 HTRWEYCAITCADNTMNDTDVPLETTECICQGGEGYRGTVNTIWNIGIPCORWDSQYPHE 305
Db 272 HTRWEYCAITCADNTMNDTDVPLETTECICQGGEGYRGTVNTIWNIGIPCORWDSQYPHE 331
QY 306 HDMPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIIPNCMDSHGDCYRGNG 365
Db 332 HDMPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIIPNCMDSHGDCYRGNG 391
QY 366 KNYMGNLSQTRSGLTCSMWKDNMEDLHRHIFWEPDASKLNEYCRNPDHAGHPWCYTG 425
Db 392 KNYMGNLSQTRSGLTCSMWKDNMEDLHRHIFWEPDASKLNEYCRNPDHAGHPWCYTG 451
QY 426 PLIPWDYCPISRCGGDTTPTIV 447
Db 452 PLIPWDYCPISRCGGDTTPTIV 473

RESULT 5
US-10-133-912-19
Sequence 19, Application US/10133912
Patent No. US20020165358A1
GENERAL INFORMATION:
APPLICANT: Kinosaki, Masahiko
Yamaguchi, Kyoji
Goto, Masaaki

Murakami, Akihiko
Ueda, Masatsugu
Higashio, Kanji
Yamashita, Yasushi
TITLE OF INVENTION: TCF MUTANT
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burgess, Ryan and Wayne
STREET: 370 Lexington Avenue, Suite 2105
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: UNITED STATES OF AMERICA
ZIP: 10017
COMPUTER READABLE FORM:
MEDIUM TYPE: 3 1/5 inch diskette
COMPUTER: PC'S LIMITED SYSTEM 200
OPERATING SYSTEM: WORD FOR WINDOWS
SOFTWARE: MICROSOFT WINDOWS 98
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/133,912
FILING DATE: 25-Apr-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/700,519
FILING DATE: 26-Aug-1996
ATTORNEY/AGENT INFORMATION:
NAME: Wayne, Milton J.
REGISTRATION NUMBER: 17,906
REFERENCE/DOCKET NUMBER: U-Wp-5240PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-683-8150
TELEFAX: 212-532-4285
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 723
TYPE: AMINO ACID
STRANDEDNESS: SINGLE
TOPOLOGY: LINEAR
SEQUENCE DESCRIPTION: SEQ ID NO: 19:
US-10-133-912-19

Query Match 97.5%; Score 2539.5; DB 9; Length 723;
Best Local Similarity 97.8%; Pred. No. 1e-196;
Matches 437; Conservative 1; Mismatches 4; Indels 5; Gaps 1;
QY 1 ERKRRTIHEFKSAKTLTIKIDPALKIKTKVNTADQCANRCKTRNKGLPFTCKAFVDFK 60
Db 32 ERKRRTIHEFKSAKTLTIKIDPALKIAATAAANTADQCANRCKTRNKGLPFTCKAFVDFK 91
QY 61 ARKQCLWFPNSMSGVKEFGHEFDLYENKDYIRNCIIIGKRSYKGTVSITKSGIKCQP 120
Db 92 ARKQCLWFPNSMSGVKEFGHEFDLYENKDYIRNCIIIGKRSYKGTVSITKSGIKCQP 151
QY 121 WSMIPHEHFLPSSYRGKDLQENYCRNPRGEGGPGWCFSTSNPEVRYEVCIDIPOCSEVEC 180
Db 152 WSMIPHEH-----SYRGKDLQENYCRNPRGEGGPGWCFSTSNPEVRYEVCIDIPOCSEVEC 206
QY 181 MTCNGESYRGLMDHTESGKICQWHDQTPHRRKFLPERYPDKGFDNDYCRNPDGQPRPWC 240
Db 207 MTCNGESYRGLMDHTESGKICQWHDQTPHRRKFLPERYPDKGFDNDYCRNPDGQPRPWC 266
QY 241 YTLDPHTRWEYCAITCADNTMNDTDVPLETTECICQGGEGYRGTVNTIWNIGIPCORWDS 300
Db 267 YTLDPHTRWEYCAITCADNTMNDTDVPLETTECICQGGEGYRGTVNTIWNIGIPCORWDS 326
QY 301 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIIPNCMDSHGDC 360
Db 327 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIIPNCMDSHGDC 386
QY 361 YRGNKKNYMGNLSQTRSGLTCSMWKDNMEDLHRHIFWEPDASKLNEYCRNPDHAGHPW 420
Db 387 YRGNKKNYMGNLSQTRSGLTCSMWKDNMEDLHRHIFWEPDASKLNEYCRNPDHAGHPW 446

Db 443 SVRWEYCNLKR-SETGGSV 462

RESULT 8

US-10-131-241-1

; Sequence 1, Application US/10131241

; Publication No. US20030012792A1

; GENERAL INFORMATION:

; APPLICANT: Holaday, John W.

; APPLICANT: Fortier, Anne H.

; TITLE OF INVENTION: Compositions and Methods for Inhibiting Endothelial Cell Proliferation

; FILE REFERENCE: 05213-0344 43170-271565

; CURRENT FILING DATE: 2002-07-22

; CURRENT APPLICATION NUMBER: US/10131,241

; PRIOR FILING DATE: 1998-05-22

; PRIOR APPLICATION NUMBER: US 09/413,049

; PRIOR FILING DATE: 1999-10-06

; PRIOR APPLICATION NUMBER: US 09/316,802

; PRIOR FILING DATE: 1999-05-21

; PRIOR APPLICATION NUMBER: US 60/086,586

; PRIOR FILING DATE: 1998-05-22

; NUMBER OF SEQ ID NOS: 65

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 1

; LENGTH: 812

; TYPE: PRT

; ORGANISM: Murinae sp.

US-10-131-241-1

Query Match 40.7%; Score 1059; DB 9; Length 812;

Best Local Similarity 43.3%; Pred. No. 2.8e-77;

Matches 191; Conservative 67; Mismatches 145; Indels 38; Gaps 12;

Qy 30 TKKVNTA--DQCANRCTRNKGLPFTCKAFVFDKARKOCLWFPFNSMSSGVKKKEGHEFD 86

Db 37 TKKQLAAGVSDCLAKCEG--TDFVCRSFQYHSKEQOCVMAENSKTSSIIIRM--RDVI 92

Qy 87 LYENKDYIRNCIIIGKRGYKGTIVTSITKSGIKQCPWSSMIPHEHSLPSSYRGKDLQENYC 146

Db 93 LFEKRVYLSCKTGIGNGYRGTMSTRTSGVACQKKGATFPHVPNTSPSTHPNEGLEENYC 152

Qy 147 RNPREGGPGWCTSNPEVRYEVCIDIPOCSEVECTNCGESYRGLMDHTESGKICQWDH 206

Db 153 RNPNDDEGPGWCTTDPKRYDYCNIPCEE--ECMYCSEGEYEGKISKTMSGLDQAWDS 211

Qy 207 QTPHRHKLPERYPDKGFDNDYCRNPDGQPRPWCYTLDPHTRWEYCAIKTCADNTMNDTD 266

Db 212 QSPHAGYIPAKFSPKLNKMYCHNPGEPRPWCFTTDTPTKRWEYCDIPRCT---TTPP 267

Qy 267 VPLETECIOGQGEYRGTVNTIWNIGIPCQRWDSQYPHEHDMTPENFKCKDLRENYCRNP 326

Db 268 PPSPTYQCLKGRGENYRGTVSVTSGKTCQRMSEQTPHRRNRTPENFPCKNLEENYCNP 327

Qy 327 DGSSESPWCTTDPNIRVGYCSQIPNCDMSHG-----ODCYRGNGKNYM 369

Db 328 DGETAPWCYTDDSQLRWEYC-EIPSCSASSPDQSDSSVPPEEQTPVVOEQYQSDGQSYR 386

Qy 370 GNLSQTRSLGSCMMDKNEDLHRHIFWE---PDASKLNENYCRNPDDDAHGPWCYTCNP 426

Db 387 GTSSTITGKKCOSWAAMP--HRHSKTPENFPDAG-LEMNYCRNPDDG-KGPWCYTDDP 442

Qy 427 LIPWDYCPISRCGEGDTTPIV 447

Db 443 SVRWEYCNLKR-SETGGSV 462

RESULT 9

US-09-788-142-1

; Sequence 1, Application US/09788142

; Patent No. US20010029246A1

; GENERAL INFORMATION:

; APPLICANT: Folkman, M. Judah

; O'Reilly, Michael

TITLE OF INVENTION: Angiostatin Fragments and Method of Use

NUMBER OF SEQUENCES: 6

CORRESPONDENCE ADDRESS:

ADDRESSEE: Jones & Askew, LLP

STREET: 191 Peachtree Street, 37th Floor

CITY: Atlanta

STATE: Georgia

COUNTRY: USA

ZIP: 30303-1769

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/788,142

FILING DATE: 16-Feb-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/866,735

FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Warren, William L.

REGISTRATION NUMBER: 36,714

REFERENCE/DOCKET NUMBER: 05940-0129

TELECOMMUNICATION INFORMATION:

TELEPHONE: (404) 818-3700

TELEFAX: (404) 818-3799

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 812 amino acids

TYPE: amino acid

STRANDEDNESS: <Unknown>

TOPOLOGY: linear

MOLECULE TYPE: protein

HYPOTHETICAL: NO

ANTI-SENSE: NO

FRAGMENT TYPE: N-terminal

ORIGINAL SOURCE:

ORGANISM: Murine

IMMEDIATE SOURCE:

CLONE: Plasmidogen

SEQUENCE DESCRIPTION: SEQ ID NO: 1:

US-09-788-142-1

Query Match 40.7%; Score 1059; DB 10; Length 812;

Best Local Similarity 43.3%; Pred. No. 2.8e-77;

Matches 191; Conservative 67; Mismatches 145; Indels 38; Gaps 12;

Qy 30 TKKVNTA--DQCANRCTRNKGLPFTCKAFVFDKARKOCLWFPFNSMSSGVKKKEGHEFD 86

Db 37 TKKQLAAGVSDCLAKCEG--TDFVCRSFQYHSKEQOCVMAENSKTSSIIIRM--RDVI 92

Qy 87 LYENKDYIRNCIIIGKRGYKGTIVTSITKSGIKQCPWSSMIPHEHSLPSSYRGKDLQENYC 146

Db 93 LFEKRVYLSCKTGIGNGYRGTMSTRTSGVACQKKGATFPHVPNTSPSTHPNEGLEENYC 152

Qy 147 RNPREGGPGWCTSNPEVRYEVCIDIPOCSEVECTNCGESYRGLMDHTESGKICQWDH 206

Db 153 RNPNDDEGPGWCTTDPKRYDYCNIPCEE--ECMYCSEGEYEGKISKTMSGLDQAWDS 211

Qy 207 QTPHRHKLPERYPDKGFDNDYCRNPDGQPRPWCYTLDPHTRWEYCAIKTCADNTMNDTD 266

Db 212 QSPHAGYIPAKFSPKLNKMYCHNPGEPRPWCFTTDTPTKRWEYCDIPRCT---TTPP 267

Qy 267 VPLETECIOGQGEYRGTVNTIWNIGIPCQRWDSQYPHEHDMTPENFKCKDLRENYCRNP 326

Db 268 PPSPTYQCLKGRGENYRGTVSVTSGKTCQRMSEQTPHRRNRTPENFPCKNLEENYCNP 327

Qy 327 DGSSESPWCTTDPNIRVGYCSQIPNCDMSHG-----ODCYRGNGKNYM 369

Db 328 DGETAPWCYTDDSQLRWEYC-EIPSCSASSPDQSDSSVPPEEQTPVVOEQYQSDGQSYR 386

QY 370 GNLSTQSLGTCMSWMDKLNEDLHRHIFWE---PDASKLNENYCRNPDDDAHGWCYTGPNP 426
Db 387 GTSSTITGKCKQSWAAMFP--HRHSTPENFPDAG--LENNYCRNPDDG--KGPWCYTTPD 442
QY 427 LIPWDYCPISRCGEGDTPTIV 447
Db 443 SVRWEYCNLKR--SETGGSVV 462

RESULT 10

US-09-761-120-1
; Sequence 1, Application US/09761120
; Patent No. US20020037847A1
; GENERAL INFORMATION:
; APPLICANT: Folkman, M. Judah
; APPLICANT: O'Reilly, Michael
; TITLE OF INVENTION: Nucleic Acids Encoding Kringle 1-5 Region Fragments of Plasminogen
; FILE REFERENCE: 05940-0151 (43171-252068)
; CURRENT APPLICATION NUMBER: US/09761,120
; CURRENT FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: 09/309,821
; PRIOR FILING DATE: 1999-05-11
; PRIOR APPLICATION NUMBER: 08/866,735
; PRIOR FILING DATE: 1997-05-30
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 812
; TYPE: PRT
; ORGANISM: Murinae gen. sp.
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Plasminogen
US-09-761-120-1

Query Match 40.7%; Score 1059; DB 10; Length 812;
Best Local Similarity 43.3%; Pred. No. 2.8e-77;
Matches 191; Conservative 67; Mismatches 145; Indels 38; Gaps 12;

QY 30 TKKVNTA---DQCANRCTRNGKLPFTCKAFVDFKARKOCLWFFPNSMSSGVKKEGHEFD 86
Db 37 TKKLAAGVSDCLAKCEG--TDFVCRSFQYHSKEQCVMAENSKTSSIRM--RDVI 92
QY 87 LYENKDYIRNCIIIGKRGYSKGVTSITKGIKQCPWSSMIPHEHSLPSSYRGKDLQENYC 146
Db 93 LFEKRVYLSCKTGICNGYRGMTSRKSGVACQKGGATFPHPVNTSPSTHPNEGLEENYC 152
QY 147 RNPGRGEGPWCFTSNPEVRYEVCDDIPQCSVECMTCNGESYRGLMDHTESGKICQDND 206
Db 153 RNPDNDEQGPWCYTTPDKRYDNCIPECEE--ECMYCSGKEYGKISKTMSGLDQAWDS 211
QY 207 QTPHRHKLPERYPDKGFDNDYCRNPDPGPRWCYTLDPHTRWEYCAIKTCADNTMNDTD 266
Db 212 QSPHAGYIPAKFPNKLKMNCHNPDGEPWPWCFTTDTTKRWEYCDIPRCI---TPPP 267
QY 267 VPLETTECIGOGEGYRGVTINWGPCQWDSQYPHEHDMTPENFKCKDLRENYCRNP 326
Db 268 PPSPTYQCLKRGENYRGVTSVTSVSGTKCQWSEQTPHRHNTPENFPCKNLEENYCRNP 327
QY 327 DGSESPWCFTTDPNIRVGYCSQIPNCDMSHG-----ODCYRGNKNYM 369
Db 328 DGETAPWCYTTSQRLWEYC--EIPSCSSASPDQSDSSVPPPEQTPVVOECYQSDGQSTR 386
QY 370 GNLSTQSLGTCMSWMDKLNEDLHRHIFWE---PDASKLNENYCRNPDDDAHGWCYTGPNP 426
Db 387 GTSSTITGKCKQSWAAMFP--HRHSTPENFPDAG--LENNYCRNPDDG--KGPWCYTTPD 442
QY 427 LIPWDYCPISRCGEGDTPTIV 447
Db 443 SVRWEYCNLKR--SETGGSVV 462

RESULT 11

US-09-873-676-81
; Sequence 81, Application US/09873676
; Patent No. US20020077289A1
; GENERAL INFORMATION:
; APPLICANT: Macdonald, Nicholas J.
; APPLICANT: Sim, Kim L.
; TITLE OF INVENTION: Angiostatin and Endostatin Binding Proteins and Methods of Use
; FILE REFERENCE: 05213-0378 (43170-259333)
; CURRENT APPLICATION NUMBER: US/09/873,676
; CURRENT FILING DATE: 2001-06-04
; PRIOR APPLICATION NUMBER: US 60/209,065
; PRIOR FILING DATE: 2000-06-02
; PRIOR APPLICATION NUMBER: US 60/289,387
; PRIOR FILING DATE: 2001-05-08
; NUMBER OF SEQ ID NOS: 123
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 81
; LENGTH: 812
; TYPE: PRT
; ORGANISM: Murinae gen. sp.
US-09-873-676-81

Query Match 40.7%; Score 1059; DB 10; Length 812;
Best Local Similarity 43.3%; Pred. No. 2.8e-77;
Matches 191; Conservative 67; Mismatches 145; Indels 38; Gaps 12;

QY 30 TKKVNTA---DQCANRCTRNGKLPFTCKAFVDFKARKOCLWFFPNSMSSGVKKEGHEFD 86
Db 37 TKKLAAGVSDCLAKCEG--TDFVCRSFQYHSKEQCVMAENSKTSSIRM--RDVI 92
QY 87 LYENKDYIRNCIIIGKRGYSKGVTSITKGIKQCPWSSMIPHEHSLPSSYRGKDLQENYC 146
Db 93 LFEKRVYLSCKTGICNGYRGMTSRKSGVACQKGGATFPHPVNTSPSTHPNEGLEENYC 152
QY 147 RNPGRGEGPWCFTSNPEVRYEVCDDIPQCSVECMTCNGESYRGLMDHTESGKICQDND 206
Db 153 RNPDNDEQGPWCYTTPDKRYDNCIPECEE--ECMYCSGKEYGKISKTMSGLDQAWDS 211
QY 207 QTPHRHKLPERYPDKGFDNDYCRNPDPGPRWCYTLDPHTRWEYCAIKTCADNTMNDTD 266
Db 212 QSPHAGYIPAKFPNKLKMNCHNPDGEPWPWCFTTDTTKRWEYCDIPRCI---TPPP 267
QY 267 VPLETTECIGOGEGYRGVTINWGPCQWDSQYPHEHDMTPENFKCKDLRENYCRNP 326
Db 268 PPSPTYQCLKRGENYRGVTSVTSVSGTKCQWSEQTPHRHNTPENFPCKNLEENYCRNP 327
QY 327 DGSESPWCFTTDPNIRVGYCSQIPNCDMSHG-----ODCYRGNKNYM 369
Db 328 DGETAPWCYTTSQRLWEYC--EIPSCSSASPDQSDSSVPPPEQTPVVOECYQSDGQSTR 386
QY 370 GNLSTQSLGTCMSWMDKLNEDLHRHIFWE---PDASKLNENYCRNPDDDAHGWCYTGPNP 426
Db 387 GTSSTITGKCKQSWAAMFP--HRHSTPENFPDAG--LENNYCRNPDDG--KGPWCYTTPD 442
QY 427 LIPWDYCPISRCGEGDTPTIV 447
Db 443 SVRWEYCNLKR--SETGGSVV 462

RESULT 12

US-09-946-893-5
; Sequence 5, Application US/09946893
; Patent No. US20020072494A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yihai
; TITLE OF INVENTION: Materials and methods relating to endothelial cell growth
; FILE REFERENCE: Mewburn
; CURRENT APPLICATION NUMBER: US/09/946,893
; CURRENT FILING DATE: 2001-09-05
; PRIOR APPLICATION NUMBER: US 60/230,893
; PRIOR FILING DATE: 2000-09-05
; NUMBER OF SEQ ID NOS: 11

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 5

; LENGTH: 569

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: Angioquiescin
US-09-946-893-5

Query Match 40.0%; Score 1042; DB 10; Length 569;

Best Local Similarity 43.6%; Pred. No. 4.3e-76;

Matches 193; Conservative 68; Mismatches 140; Indels 42; Gaps 13;

Qy 30 TKK---VNTADQCANRCTRNGKLPFTCKAFVDFKARKQCLWFFPNFMSGSGVKKFEGHEFD 86

Db 37 TKKQLGAGSIEECAAACEDE--EFTCRAFOYHSKEQOCVMAENRKSIIIRM--RDVV 92

Qy 87 LYENKDYIRNCLIGKRSYKGTVSITKSGIKQCPWSSMIPHEHSFLPSSYRGKDLQENYC 146

Db 93 LFEKKVYLSECKTGNGKNYRGTMSKTKNGITCQKWSSTSPHRPRFSPATHPSEGLEENYC 152

Qy 147 RNPGEEGGPMWCFSTNPEVRYEVEDIPQSEVECMTCNGESYRGLMDHTESGKICQRDWH 206

Db 153 RNPNDPOGPMWCIYTDPEKRYDYCDILECEE--ECMHCSGENYDGKISKTMSGLECQAWDS 211

Qy 207 QTPHRHKFLPERYPDKGFDNNCRNPDGQPRWCYTLDPHTRWEYCAIKTCADNTMNDTD 266

Db 212 QSPHAGHYIPSKFKNLKNKYNCRNPDRELPRWCFTTDPNKRKWLCDIPRCT-----TP 265

Qy 267 VPLE--TTECTOGOGEGYRGVTNINWIGIPCORWDSQYPHEHDMTPENFKCKDLRENVCYR 324

Db 266 PPSGGPTYQCLKGTGENTYRGNVAVTVSGHTCOHWSAQTPTHNRTPENFPCKNLNDENVCYR 325

Qy 325 NPDGSESPWCFTTDPNIRVGYCSQIPNCDMSH-----GDCYRGNGKN 367

Db 326 NPDGRAPWCHTTSQVRWEYC--KIPSCDSSPVSTEQLAPTAPPETLPVQDCYHGDQS 384

Qy 368 YMGNLSTQSRGLTCSMDKNMEDLHRHFW---PDASKLNEYCRNPDHDAHGPWCYTG 424

Db 385 YRGTSSTTTTKKCKQSSWMT--HRHOKTPENYPNAG--LTMNYCRNPAD--KGPWCFTT 440

Qy 425 NPLIPWDYCPISRCEDTPTIV 447

Db 441 DPSVRWEYCNLKKCSG--TEASVV 462

RESULT 13

US-09-946-893-8

; Sequence 8, Application US/09946893

; Patent No. US20020072494A1

; GENERAL INFORMATION:

; APPLICANT: Cao, Yihai

; TITLE OF INVENTION: Materials and methods relating to endothelial cell growth

; FILE REFERENCE: Mewburn

; CURRENT APPLICATION NUMBER: US/09/946,893

; PRIOR FILING DATE: 2001-09-05

; PRIOR APPLICATION NUMBER: US 60/230,893

; NUMBER OF SEQ ID NOS: 11

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 8

; LENGTH: 571

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: Angioquiescin
US-09-946-893-8

Query Match 40.0%; Score 1042; DB 10; Length 571;

Best Local Similarity 43.6%; Pred. No. 4.3e-76;

Matches 193; Conservative 68; Mismatches 140; Indels 42; Gaps 13;

QY 368 YMGNLSQTRSGLTCSMWKDNMEDLHRHIFWE---PDASKLNENYCRNPDDDAHGWCYTG 424
Db 366 YRGTSSTTTTGGKQCSWSSMTP--HRHOKTPENYPNAG-LTMNYCRNPAD-KGPWCFTT 421
QY 425 NPLIPWDYCPISRCBGDTTPTIV 447
Db 422 DPSVRWEYCNLKKCSG-TEASVV 443

Search completed: April 25, 2003, 15:49:03
Job time : 25 secs

QY 207 QTPHRHKLPERYPDKGFDNDYCRNPDPQRPWCYTILDPHTRWEYCAIKTCAADTMNDTD 266
Db 212 QSPHAGYIPSKFPPKNLKNKYNCRNPDRPWCFTTDPNKRWELCDIPRCT-----TP 265
QY 267 VPLE--TTECIQOGGEGYRGVNTIWIWGPCQWDSQVPEHDMTPENFKCKDLRENYCR 324
Db 266 PPSSGPTYQCLKGKGTGENYRGVAVTVSGHTCQHSAGTPTHNRTPTENFPCKNLDENYCR 325
QY 325 NPDGSESPWCFTTDPNIRVGYCQIPNCDMSH-----GQDCYRGNKGN 367
Db 326 NPDGKRAPWCHTNSQVRWEYC-KIPSCDSSPVSTEQALAPTAPPELTVPVQDCYHGDGQS 384
QY 368 YMGNLSQTRSGLTCSMWKDNMEDLHRHIFWE---PDASKLNENYCRNPDDDAHGWCYTG 424
Db 385 YRGTSSTTTTGGKQCSWSSMTP--HRHOKTPENYPNAG-LTMNYCRNPAD-KGPWCFTT 440
QY 425 NPLIPWDYCPISRCBGDTTPTIV 447
Db 441 DPSVRWEYCNLKKCSG-TEASVV 462

RESULT 15
US-09-967-386-1
; Sequence 1, Application US/09967386
; Patent No. US20020159992A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Henkin, Jack
; TITLE OF INVENTION: ANTIANGIOGENIC POLYPEPTIDES AND METHODS
; TITLE OF INVENTION: FOR INHIBITING ANGIOGENESIS
; FILE REFERENCE: 6738 US, 02
; CURRENT APPLICATION NUMBER: US/09/967,386
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: US 60/236,550
; PRIOR FILING DATE: 2000-09-29
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 791
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-967-386-1

Query Match 40.0%; Score 1042; DB 9; Length 791;
Best Local Similarity 43.6%; Pred. No. 6,4e-76;
Matches 193; Conservative 68; Mismatches 140; Indels 42; Gaps 13;

QY 30 TKK---VNTADOCANRCTRNLGLPFTCKAFVFDKARKQCLWFPFNSMSGVKKKEFGHEFD 86
Db 18 TKKQLGAGSIECAKCEDE--EFTCRAFYHSEKQCVIMAENRKSIIIRM--RDVV 73
QY 87 LYENKDIYRNCIIIGKRSYKGTIVTSKGIKQCPWSSMIPHEHSELPSSYRGKDLQENYC 146
Db 74 LFEKKVYLSECKTGNGKNGYGTMTKNGITCKQWSTSPHRPRFSPATHPSEGLEENYC 133
QY 147 RNPREGGEGPWCFTSNPEVRYEVCIDIPQCSVECEMTCNGESYRGLMDHTESGKICQWHDH 206
Db 134 RNPNDPQGPWCYTTPDEKRYDYCDILECEE-ECMHCSGENYDGKISKTMSGLECCQAWDS 192
QY 207 QTPHRHKLPERYPDKGFDNDYCRNPDPQRPWCYTILDPHTRWEYCAIKTCAADTMNDTD 266
Db 193 QSPHAGYIPSKFPPKNLKNKYNCRNPDRPWCFTTDPNKRWELCDIPRCT-----TP 246
QY 267 VPLE--TTECIQOGGEGYRGVNTIWIWGPCQWDSQVPEHDMTPENFKCKDLRENYCR 324
Db 247 PPSSGPTYQCLKGKGTGENYRGVAVTVSGHTCQHSAGTPTHNRTPTENFPCKNLDENYCR 306
QY 325 NPDGSESPWCFTTDPNIRVGYCQIPNCDMSH-----GQDCYRGNKGN 367
Db 307 NPDGKRAPWCHTNSQVRWEYC-KIPSCDSSPVSTEQALAPTAPPELTVPVQDCYHGDGQS 365

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 25, 2003, 15:40:48 ; Search time 145 Seconds
(without alignments)
1987.557 Million cell updates/sec

Title: US-09-674-377B-1

Perfect score: 2604

Sequence: 1 ERKRRNTIHEFKSAKTTLI.....IPWDYCPISRCGGTTPTIV 447

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 4569144 seqs, 644733110 residues

Total number of hits satisfying chosen parameters: 4569144

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending_Patents_AA_Main.*

```

1: /cgn2_6/ptodata/1/paa/PCRUS_COMB.pep.*
2: /cgn2_6/ptodata/1/paa/US06_COMB.pep.*
3: /cgn2_6/ptodata/1/paa/US07_COMB.pep.*
4: /cgn2_6/ptodata/1/paa/US080_COMB.pep.*
5: /cgn2_6/ptodata/1/paa/US081_COMB.pep.*
6: /cgn2_6/ptodata/1/paa/US082_COMB.pep.*
7: /cgn2_6/ptodata/1/paa/US083_COMB.pep.*
8: /cgn2_6/ptodata/1/paa/US084_COMB.pep.*
9: /cgn2_6/ptodata/1/paa/US085_COMB.pep.*
10: /cgn2_6/ptodata/1/paa/US086_COMB.pep.*
11: /cgn2_6/ptodata/1/paa/US087_COMB.pep.*
12: /cgn2_6/ptodata/1/paa/US088_COMB.pep.*
13: /cgn2_6/ptodata/1/paa/US089_COMB.pep.*
14: /cgn2_6/ptodata/1/paa/US090_COMB.pep.*
15: /cgn2_6/ptodata/1/paa/US091_COMB.pep.*
16: /cgn2_6/ptodata/1/paa/US092_COMB.pep.*
17: /cgn2_6/ptodata/1/paa/US093_COMB.pep.*
18: /cgn2_6/ptodata/1/paa/US094_COMB.pep.*
19: /cgn2_6/ptodata/1/paa/US095_COMB.pep.*
20: /cgn2_6/ptodata/1/paa/US096_COMB.pep.*
21: /cgn2_6/ptodata/1/paa/US097_COMB.pep.*
22: /cgn2_6/ptodata/1/paa/US098_COMB.pep.*
23: /cgn2_6/ptodata/1/paa/US099_COMB.pep.*
24: /cgn2_6/ptodata/1/paa/US100_COMB.pep.*
25: /cgn2_6/ptodata/1/paa/US101_COMB.pep.*
26: /cgn2_6/ptodata/1/paa/US102_COMB.pep.*
27: /cgn2_6/ptodata/1/paa/US60_COMB.pep.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2604	100.0	447	20	US-09-674-377B-1
2	2601	99.9	447	23	US-09-951-629-1
3	2601	99.9	447	24	US-10-081-309-2
4	2601	99.9	697	5	US-08-167-641A-44
5	2601	99.9	697	8	US-08-458-217-50
6	2601	99.9	697	8	US-08-460-890-44

```

7 2601 99.9 697 8 US-08-460-971-50
8 2601 99.9 728 3 US-07-705-741-1
9 2601 99.9 728 3 US-07-712-932-3
10 2601 99.9 728 3 US-07-983-366-4
11 2601 99.9 728 6 US-08-242-802-2
12 2601 99.9 728 6 US-08-292-160B-1
13 2601 99.9 728 10 US-08-605-221-2
14 2601 99.9 728 13 US-08-909-515-2
15 2601 99.9 728 13 US-08-930-999-1
16 2601 99.9 728 18 US-09-423-516-2
17 2601 99.9 728 21 US-09-791-537-140648
18 2592 99.5 728 13 US-08-963-643-1
19 2566.5 98.6 442 20 US-09-674-377B-2
20 2565 98.5 728 21 US-09-791-537-36146
21 2565 98.5 728 21 US-09-791-537-140804
22 2565 98.5 728 27 US-60-382-445-2889
23 2563.5 98.4 487 20 US-09-601-040A-27
24 2563.5 98.4 717 20 US-09-601-040-6
25 2563.5 98.4 717 20 US-09-601-040-8
26 2563.5 98.4 717 20 US-09-601-040A-6
27 2563.5 98.4 717 20 US-09-601-040A-8
28 2563.5 98.4 723 3 US-07-712-932-5
29 2563.5 98.4 723 6 US-08-267-837-4
30 2563.5 98.4 723 6 US-08-290-937-1
31 2563.5 98.4 723 6 US-08-290-937-2
32 2563.5 98.4 723 10 US-08-605-221-4
33 2563.5 98.4 723 20 US-09-600-991-18
34 2563.5 98.4 723 20 US-09-601-040-10
35 2563.5 98.4 723 20 US-09-601-040A-10
36 2563.5 98.4 723 21 US-09-791-537-140107
37 2563.5 98.4 729 20 US-09-601-040-2
38 2563.5 98.4 729 20 US-09-601-040-4
39 2563.5 98.4 729 20 US-09-601-040A-2
40 2563.5 98.4 729 20 US-09-601-040A-4
41 2557.5 98.2 723 3 US-07-705-741-2
42 2557.5 98.2 723 6 US-08-290-937-3
43 2557.5 98.2 723 6 US-08-292-160B-2
44 2545 97.6 836 27 US-60-230-435-1887
45 2541.5 97.6 723 25 US-10-133-912-18

```

ALIGNMENTS

```

RESULT 1
US-09-674-377B-1
; Sequence 1, Application US/09674377B
; GENERAL INFORMATION:
; APPLICANT: Nakamura, Toshikazu
; TITLE OF INVENTION: NEOVASCULARIZATION INHIBITORS
; FILE REFERENCE: Q61434
; CURRENT APPLICATION NUMBER: US/09/674,377B
; CURRENT FILING DATE: 2000-10-30
; PRIOR APPLICATION NUMBER: PCT/JF99/01834
; PRIOR FILING DATE: 1999-04-06
; PRIOR APPLICATION NUMBER: JP 1998-134681
; PRIOR FILING DATE: 1998-04-28
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 447
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (1)..(1)
; OTHER INFORMATION: pyroglutamate
; FEATURE:
; NAME/KEY: CHAIN
; LOCATION: (1)..(447)
; OTHER INFORMATION: N-terminal region of alpha-chain in HGF (PyrGlu32-val478/HGF)
US-09-674-377B-1

```

Sequence 50, Appl
Sequence 1, Appli
Sequence 3, Appli
Sequence 4, Appli
Sequence 2, Appli
Sequence 1, Appli
Sequence 2, Appli
Sequence 2, Appli
Sequence 2, Appli
Sequence 140648,
Sequence 1, Appli
Sequence 2, Appli
Sequence 36146, A
Sequence 140804,
Sequence 2889, Ap
Sequence 27, Appl
Sequence 6, Appli
Sequence 8, Appli
Sequence 6, Appli
Sequence 8, Appli
Sequence 5, Appli
Sequence 4, Appli
Sequence 1, Appli
Sequence 2, Appli
Sequence 4, Appli
Sequence 10, Appl
Sequence 10, Appl
Sequence 140107,
Sequence 2, Appli
Sequence 4, Appli
Sequence 2, Appli
Sequence 2, Appli
Sequence 2, Appli
Sequence 3, Appli
Sequence 1887, Ap
Sequence 18, Appl

```
Query Match      100.0%; Score 2604; DB 20; Length 447;
Best Local Similarity 100.0%; Pred. No. 2.9e-230;
Matches 447; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ERKRRNTIHEFKSAKTTLIKIDPALKIKTKKVNADQCANRCTRNKGLPFTCKAFVFDK 60
Db 1 ERKRRNTIHEFKSAKTTLIKIDPALKIKTKKVNADQCANRCTRNKGLPFTCKAFVFDK 60

Qy 61 ARKQCLWFFPNSMSGVKKFEGHFDLYENKDYIRNCIIIGKRSYKGTVSTIKSGIKCQP 120
Db 61 ARKQCLWFFPNSMSGVKKFEGHFDLYENKDYIRNCIIIGKRSYKGTVSTIKSGIKCQP 120

Qy 121 WSSMIPHEHSFLPSSYRGKDLQENYCRNPRGEGGPGWFTSNPEVRYEVCDDIPQSEVEC 180
Db 121 WSSMIPHEHSFLPSSYRGKDLQENYCRNPRGEGGPGWFTSNPEVRYEVCDDIPQSEVEC 180

Qy 181 MTCNGESYRGLMDHTESGKICQWRDHTPHRHKFLPERYDPKGFDDNYCRNPDGQPRWC 240
Db 181 MTCNGESYRGLMDHTESGKICQWRDHTPHRHKFLPERYDPKGFDDNYCRNPDGQPRWC 240

Qy 241 YTLDPHTRWEYCAIKTCADNTMNDTVPLETTTCIOGEGYRGTVNTIWNIPQWRDS 300
Db 241 YTLDPHTRWEYCAIKTCADNTMNDTVPLETTTCIOGEGYRGTVNTIWNIPQWRDS 300

Qy 301 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 360
Db 301 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 360

Qy 361 YRGNKNGMGNLSQTRSGLTCSMWDKNMEDLHRHIFWEPDASKLNENYCRNPDGQPRWC 240
Db 361 YRGNKNGMGNLSQTRSGLTCSMWDKNMEDLHRHIFWEPDASKLNENYCRNPDGQPRWC 240

Qy 421 CYTGNPLIPWDYCPISRCGDTTPTIV 447
Db 421 CYTGNPLIPWDYCPISRCGDTTPTIV 447

RESULT 2
US-09-951-629-1
; Sequence 1, Application US/09951629
; GENERAL INFORMATION:
; APPLICANT: NAKAMURA, Toshikazu
; TITLE OF INVENTION: Anti-Cancer Agent
; FILE REFERENCE: 4373-0101P
; CURRENT APPLICATION NUMBER: US/09/951,629
; CURRENT FILING DATE: 2001-09-14
; PRIOR APPLICATION NUMBER: JP 300728
; PRIOR FILING DATE: 1995-10-24
; PRIOR APPLICATION NUMBER: PCT/JP96/03105
; PRIOR FILING DATE: 1996-10-23
; PRIOR APPLICATION NUMBER: US 09/471,032
; PRIOR FILING DATE: 1999-12-23
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 1
; LENGTH: 447
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-951-629-1

Query Match      99.9%; Score 2601; DB 23; Length 447;
Best Local Similarity 99.8%; Pred. No. 5.5e-230;
Matches 446; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ERKRRNTIHEFKSAKTTLIKIDPALKIKTKKVNADQCANRCTRNKGLPFTCKAFVFDK 60
Db 1 ERKRRNTIHEFKSAKTTLIKIDPALKIKTKKVNADQCANRCTRNKGLPFTCKAFVFDK 60

Qy 61 ARKQCLWFFPNSMSGVKKFEGHFDLYENKDYIRNCIIIGKRSYKGTVSTIKSGIKCQP 120
Db 61 ARKQCLWFFPNSMSGVKKFEGHFDLYENKDYIRNCIIIGKRSYKGTVSTIKSGIKCQP 120

Qy 121 WSSMIPHEHSFLPSSYRGKDLQENYCRNPRGEGGPGWFTSNPEVRYEVCDDIPQSEVEC 180
Db 121 WSSMIPHEHSFLPSSYRGKDLQENYCRNPRGEGGPGWFTSNPEVRYEVCDDIPQSEVEC 180
```

```
Db 121 WSSMIPHEHSFLPSSYRGKDLQENYCRNPRGEGGPGWFTSNPEVRYEVCDDIPQSEVEC 180
Qy 181 MTCNGESYRGLMDHTESGKICQWRDHTPHRHKFLPERYDPKGFDDNYCRNPDGQPRWC 240
Db 181 MTCNGESYRGLMDHTESGKICQWRDHTPHRHKFLPERYDPKGFDDNYCRNPDGQPRWC 240
Qy 241 YTLDPHTRWEYCAIKTCADNTMNDTVPLETTTCIOGEGYRGTVNTIWNIPQWRDS 300
Db 241 YTLDPHTRWEYCAIKTCADNTMNDTVPLETTTCIOGEGYRGTVNTIWNIPQWRDS 300
Qy 301 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 360
Db 301 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 360
Qy 361 YRGNKNGMGNLSQTRSGLTCSMWDKNMEDLHRHIFWEPDASKLNENYCRNPDGQPRWC 240
Db 361 YRGNKNGMGNLSQTRSGLTCSMWDKNMEDLHRHIFWEPDASKLNENYCRNPDGQPRWC 240
Qy 421 CYTGNPLIPWDYCPISRCGDTTPTIV 447
Db 421 CYTGNPLIPWDYCPISRCGDTTPTIV 447

RESULT 3
US-10-081-309-2
; Sequence 2, Application US/10081309
; GENERAL INFORMATION:
; APPLICANT: Hoffmann-La Roche Inc.
; TITLE OF INVENTION: PEG Conjugates of NK4
; FILE REFERENCE: 20859
; CURRENT APPLICATION NUMBER: US/10/081,309
; CURRENT FILING DATE: 2002-06-04
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 2
; LENGTH: 447
; TYPE: PRT
; ORGANISM: homosapiens
US-10-081-309-2

Query Match      99.9%; Score 2601; DB 24; Length 447;
Best Local Similarity 99.8%; Pred. No. 5.5e-230;
Matches 446; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ERKRRNTIHEFKSAKTTLIKIDPALKIKTKKVNADQCANRCTRNKGLPFTCKAFVFDK 60
Db 1 ERKRRNTIHEFKSAKTTLIKIDPALKIKTKKVNADQCANRCTRNKGLPFTCKAFVFDK 60

Qy 61 ARKQCLWFFPNSMSGVKKFEGHFDLYENKDYIRNCIIIGKRSYKGTVSTIKSGIKCQP 120
Db 61 ARKQCLWFFPNSMSGVKKFEGHFDLYENKDYIRNCIIIGKRSYKGTVSTIKSGIKCQP 120

Qy 121 WSSMIPHEHSFLPSSYRGKDLQENYCRNPRGEGGPGWFTSNPEVRYEVCDDIPQSEVEC 180
Db 121 WSSMIPHEHSFLPSSYRGKDLQENYCRNPRGEGGPGWFTSNPEVRYEVCDDIPQSEVEC 180

Qy 181 MTCNGESYRGLMDHTESGKICQWRDHTPHRHKFLPERYDPKGFDDNYCRNPDGQPRWC 240
Db 181 MTCNGESYRGLMDHTESGKICQWRDHTPHRHKFLPERYDPKGFDDNYCRNPDGQPRWC 240

Qy 241 YTLDPHTRWEYCAIKTCADNTMNDTVPLETTTCIOGEGYRGTVNTIWNIPQWRDS 300
Db 241 YTLDPHTRWEYCAIKTCADNTMNDTVPLETTTCIOGEGYRGTVNTIWNIPQWRDS 300

Qy 301 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 360
Db 301 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 360

Qy 361 YRGNKNGMGNLSQTRSGLTCSMWDKNMEDLHRHIFWEPDASKLNENYCRNPDGQPRWC 240
Db 361 YRGNKNGMGNLSQTRSGLTCSMWDKNMEDLHRHIFWEPDASKLNENYCRNPDGQPRWC 240
```

QY 421 CYTCNPLIPWDYCPISRCGDTTPTIV 447
 Db 421 CYTCNPLIPWDYCPISRCGDTTPTIV 447

RESULT 4

US-08-167-641A-44
 ; Sequence 44, Application US/08167641A
 ; GENERAL INFORMATION:
 ; APPLICANT: Woo, Savio L.C.
 ; APPLICANT: Smith, Louis C.
 ; APPLICANT: Cristiano, Richard J.
 ; APPLICANT: Gottchalk, Stephen
 ; TITLE OF INVENTION: NUCLEIC ACID TRANSPORTER
 ; TITLE OF INVENTION: SYSTEMS AND METHODS OF
 ; TITLE OF INVENTION: USE
 ; NUMBER OF SEQUENCES: 51
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: LYON & LYON
 ; STREET: 611 West Sixth St.
 ; CITY: Los Angeles
 ; STATE: California
 ; COUNTRY: U.S.A.
 ; ZIP: 90017
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
 ; MEDIUM TYPE: storage
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: IBM MS-DOS (Ver. 5.0)
 ; SOFTWARE: WordPerfect (Ver. 5.1)
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/167,641A
 ; FILING DATE: December 14, 1993
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA: including application
 ; PRIOR APPLICATION DATA: described below:
 ; APPLICATION NUMBER: 07/855,389
 ; FILING DATE: March 20, 1992
 ; APPLICATION NUMBER: PCT/US93/02725
 ; FILING DATE: March 19, 1993
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Knight, Matthew W.
 ; REGISTRATION NUMBER: 36,846
 ; REFERENCE/DOCKET NUMBER: 205/012
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (213) 489-1600
 ; TELEFAX: (213) 955-0440
 ; TELEX: 67-3510
 ; INFORMATION FOR SEQ ID NO: 44:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 697 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-167-641A-44

Query Match 99.9%; Score 2601; DB 5; Length 697;
 Best Local Similarity 99.8%; Pred. No. 1e-229;
 Matches 446; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ERKRRTTHFEKSAKTLIKIDPALKIKTKVNTADQCANRCTRNGKGLPTCKAFVFDK 60
 Db 1 QKRKRRTTHFEKSAKTLIKIDPALKIKTKVNTADQCANRCTRNGKGLPTCKAFVFDK 60
 QY 61 ARKQCLWFPNMSGSKVKEGHEFDLYENKDYTRNCIIIGKRSYKGTVSTKSGIKCQP 120
 Db 61 ARKQCLWFPNMSGSKVKEGHEFDLYENKDYTRNCIIIGKRSYKGTVSTKSGIKCQP 120
 QY 121 WSSMIPHEHFLPSSYRGKDLQENYCNPRGEGGPMCFSTSNPEVRYEVCIDIQCSEVEC 180
 Db 121 WSSMIPHEHFLPSSYRGKDLQENYCNPRGEGGPMCFSTSNPEVRYEVCIDIQCSEVEC 180

QY 181 MTCNGESYRGLMDHTESGKICQRWDHOTPHRKELPERYPDKGFDNDYCRNPDGQPREWC 240
 Db 181 MTCNGESYRGLMDHTESGKICQRWDHOTPHRKELPERYPDKGFDNDYCRNPDGQPREWC 240
 QY 241 YLDPHTRWEYCAIKTCADNTMNDTVPLETTTCIOGQGEYRGTVNTIWNIGIFCQRWDS 300
 Db 241 YLDPHTRWEYCAIKTCADNTMNDTVPLETTTCIOGQGEYRGTVNTIWNIGIFCQRWDS 300
 QY 301 QYPHEHDMTPENFKCKDLRENYCNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGDC 360
 Db 301 QYPHEHDMTPENFKCKDLRENYCNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGDC 360
 QY 361 YRGNGKNYMGNSLQTRSGLTCSMDKKNMEDLHRIHFWEPDASKLNNYCRNPDGDAHGWP 420
 Db 361 YRGNGKNYMGNSLQTRSGLTCSMDKKNMEDLHRIHFWEPDASKLNNYCRNPDGDAHGWP 420
 QY 421 CYTCNPLIPWDYCPISRCGDTTPTIV 447
 Db 421 CYTCNPLIPWDYCPISRCGDTTPTIV 447

RESULT 5

US-08-458-217-50
 ; Sequence 50, Application US/08458217
 ; GENERAL INFORMATION:
 ; APPLICANT: Woo, Savio L.C.
 ; APPLICANT: Smith, Louis C.
 ; APPLICANT: Cristiano, Richard J.
 ; APPLICANT: Gottchalk, Stephen
 ; TITLE OF INVENTION: NUCLEIC ACID TRANSPORTER SYSTEMS AND
 ; TITLE OF INVENTION: METHODS OF USE
 ; NUMBER OF SEQUENCES: 65
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Lyon & Lyon
 ; STREET: 633 West Fifth Street
 ; CITY: Suite 4700
 ; STATE: California
 ; COUNTRY: U.S.A.
 ; ZIP: 90071-2066
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
 ; MEDIUM TYPE: storage
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: IBM P.C. DOS 5.0
 ; SOFTWARE: FastSeq for Windows 2.0
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/458,217
 ; FILING DATE: June 5, 1995
 ; CLASSIFICATION: 514
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/167,641
 ; FILING DATE: December 14, 1993
 ; APPLICATION NUMBER: 07/855,389
 ; FILING DATE: March 20, 1992
 ; APPLICATION NUMBER: PCT/US93/02725
 ; FILING DATE: March 19, 1993
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Warburg, Richard J.
 ; REGISTRATION NUMBER: 32,327
 ; REFERENCE/DOCKET NUMBER: 212/065
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (213) 489-1600
 ; TELEFAX: (213) 955-0440
 ; TELEX: 67-3510
 ; INFORMATION FOR SEQ ID NO: 50:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 697 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: peptide

US-08-458-217-50

Query Match 99.9%; Score 2601; DB 8; Length 697;
Best Local Similarity 99.8%; Pred. No. 1e-229;
Matches 446; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ERKRNTIHEFKSAKTLIKIDPALKIKTKVNTADOCANRTRNKGLPFTCKAFVFDK 60
DB 1 QKRRNTIHEFKSAKTLIKIDPALKIKTKVNTADOCANRTRNKGLPFTCKAFVFDK 60

QY 61 ARKQCLWFPNMSGGVKKKEGHEFDLYENKDYIRNCIIIGKRSYKGTVSITKSGIKCQP 120
DB 61 ARKQCLWFPNMSGGVKKKEGHEFDLYENKDYIRNCIIIGKRSYKGTVSITKSGIKCQP 120

QY 121 WSSMIPHEHSELPSSYRGKDLQENYCRNPRGEGGWCFTSNPEVRYEVCIDIQCSEVEC 180
DB 121 WSSMIPHEHSELPSSYRGKDLQENYCRNPRGEGGWCFTSNPEVRYEVCIDIQCSEVEC 180

QY 181 MTCNGESYRGLMDHTESGKICQRDHDTPHRHKFLPERYPDKGFDNNYCRNPDGQPRPWC 240
DB 181 MTCNGESYRGLMDHTESGKICQRDHDTPHRHKFLPERYPDKGFDNNYCRNPDGQPRPWC 240

QY 241 YLDPHTRWYCAIKTCADNTMNDTDVPLETTECIGOGEGYRGTVNTIWNIGPCQRWDS 300
DB 241 YLDPHTRWYCAIKTCADNTMNDTDVPLETTECIGOGEGYRGTVNTIWNIGPCQRWDS 300

QY 301 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGDC 360
DB 301 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGDC 360

QY 361 YRGNGKNYMGNLSTQTRSGLTCSMDKNMEDLHRHIFWEPDASKLNNYCRNPDDDAHGPW 420
DB 361 YRGNGKNYMGNLSTQTRSGLTCSMDKNMEDLHRHIFWEPDASKLNNYCRNPDDDAHGPW 420

QY 421 CYTGNPLIPWDYCPISRCGEGDTTPTIV 447
DB 421 CYTGNPLIPWDYCPISRCGEGDTTPTIV 447

RESULT 6

US-08-460-890-44

Sequence 44, Application US/08460890
GENERAL INFORMATION:
APPLICANT: Woo, Savio L.C.
APPLICANT: Smith, Louis C.
APPLICANT: Cristiano, Richard J.
TITLE OF INVENTION: NUCLEIC ACID TRANSPORTER
TITLE OF INVENTION: SYSTEMS AND METHODS OF
TITLE OF INVENTION: USE
NUMBER OF SEQUENCES: 51
CORRESPONDENCE ADDRESS:
ADDRESSEE: LYON & LYON
STREET: 633 West Fifth St.
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071

COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
MEDIUM TYPE: storage
COMPUTER: IBM PC compatible
OPERATING SYSTEM: IBM MS-DOS (Ver. 6.22)
SOFTWARE: Wordperfect (Ver. 5.1)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/460,890
FILING DATE: June 5, 1995
CLASSIFICATION: 435

PRIOR APPLICATION DATA:
PRIOR APPLICATION DATA: including application
PRIOR APPLICATION DATA: described below:
APPLICATION NUMBER: 08/167,641
FILING DATE: December 14, 1993

APPLICATION NUMBER: 07/855,389
FILING DATE: March 20, 1992
APPLICATION NUMBER: PCT/US93/02725
FILING DATE: March 19, 1993
ATTORNEY/AGENT INFORMATION:
NAME: Knight, Matthew W.
REGISTRATION NUMBER: 36,846
REFERENCE/DOCKET NUMBER: 212/066
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 44:
SEQUENCE CHARACTERISTICS:
LENGTH: 697 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-460-890-44

Query Match 99.9%; Score 2601; DB 8; Length 697;
Best Local Similarity 99.8%; Pred. No. 1e-229;
Matches 446; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ERKRNTIHEFKSAKTLIKIDPALKIKTKVNTADOCANRTRNKGLPFTCKAFVFDK 60
DB 1 QKRRNTIHEFKSAKTLIKIDPALKIKTKVNTADOCANRTRNKGLPFTCKAFVFDK 60

QY 61 ARKQCLWFPNMSGGVKKKEGHEFDLYENKDYIRNCIIIGKRSYKGTVSITKSGIKCQP 120
DB 61 ARKQCLWFPNMSGGVKKKEGHEFDLYENKDYIRNCIIIGKRSYKGTVSITKSGIKCQP 120

QY 121 WSSMIPHEHSELPSSYRGKDLQENYCRNPRGEGGWCFTSNPEVRYEVCIDIQCSEVEC 180
DB 121 WSSMIPHEHSELPSSYRGKDLQENYCRNPRGEGGWCFTSNPEVRYEVCIDIQCSEVEC 180

QY 181 MTCNGESYRGLMDHTESGKICQRDHDTPHRHKFLPERYPDKGFDNNYCRNPDGQPRPWC 240
DB 181 MTCNGESYRGLMDHTESGKICQRDHDTPHRHKFLPERYPDKGFDNNYCRNPDGQPRPWC 240

QY 241 YLDPHTRWYCAIKTCADNTMNDTDVPLETTECIGOGEGYRGTVNTIWNIGPCQRWDS 300
DB 241 YLDPHTRWYCAIKTCADNTMNDTDVPLETTECIGOGEGYRGTVNTIWNIGPCQRWDS 300

QY 301 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGDC 360
DB 301 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGDC 360

QY 361 YRGNGKNYMGNLSTQTRSGLTCSMDKNMEDLHRHIFWEPDASKLNNYCRNPDDDAHGPW 420
DB 361 YRGNGKNYMGNLSTQTRSGLTCSMDKNMEDLHRHIFWEPDASKLNNYCRNPDDDAHGPW 420

QY 421 CYTGNPLIPWDYCPISRCGEGDTTPTIV 447
DB 421 CYTGNPLIPWDYCPISRCGEGDTTPTIV 447

RESULT 7

US-08-460-971-50

Sequence 50, Application US/08460971
GENERAL INFORMATION:
APPLICANT: Woo, Savio L.C.
APPLICANT: Smith, Louis C.
APPLICANT: Cristiano, Richard J.
TITLE OF INVENTION: NUCLEIC ACID TRANSPORTER SYSTEMS AND
TITLE OF INVENTION: METHODS OF USE
NUMBER OF SEQUENCES: 65
CORRESPONDENCE ADDRESS:
ADDRESSEE: LYON & LYON
STREET: 633 West Fifth Street
STREET: Suite 4700

CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071-2066
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
MEDIUM TYPE: storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: FASTSEQ for Windows 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/460,971
FILING DATE: June 5, 1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/167,641
FILING DATE: December 14, 1993
APPLICATION NUMBER: 07/855,389
FILING DATE: March 20, 1992
APPLICATION NUMBER: PCT/US93/02725
FILING DATE: March 19, 1993
ATTORNEY/AGENT INFORMATION:
NAME: Warburg, Richard J.
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 212/063
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 50:
SEQUENCE CHARACTERISTICS:
LENGTH: 697 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-460-971-50

Query Match 99.9%; Score 2601; DB 8; Length 697;
Best Local Similarity 99.8%; Pred. No. 1e-229;
Matches 446; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ERKRNTTIEFKSAKTLTIKIDPALKIKTKKVNADQCANRCTRNKGLPTCKAFVFDK 60
Db :|||||
Qy 1 QKRRNTTIEFKSAKTLTIKIDPALKIKTKKVNADQCANRCTRNKGLPTCKAFVFDK 60
Db :|||||
Qy 61 ARKQCLWPFPPNSMSGVKKFEGHFDLYENKDYIRNCIIIGKRSYKGTVSTKSGIKQCP 120
Db :|||||
Qy 61 ARKQCLWPFPPNSMSGVKKFEGHFDLYENKDYIRNCIIIGKRSYKGTVSTKSGIKQCP 120
Db :|||||
Qy 121 WSSMIPHEHSLPSSYRGKDLQENYCRNPRGEGPWCFTSNPEVRYEVCIDIPOQSEVEC 180
Db :|||||
Qy 181 MTCNGESYRGLMDHTESGKICQWRDHTPHRHKFLPERYPDKGDDNYCRNPDGQPRWC 240
Db :|||||
Qy 181 MTCNGESYRGLMDHTESGKICQWRDHTPHRHKFLPERYPDKGDDNYCRNPDGQPRWC 240
Db :|||||
Qy 241 YLDPHTRWECALTKCADNTMNDTVDPLETTECIGOGEGYRGVTNINWGIQCORWDS 300
Db :|||||
Qy 241 YLDPHTRWECALTKCADNTMNDTVDPLETTECIGOGEGYRGVTNINWGIQCORWDS 300
Db :|||||
Qy 301 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 360
Db :|||||
Qy 301 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 360
Db :|||||
Qy 361 YRGNKNTMGNLSQTRSLGTCMDKNMEDLHRHIFWEPDASKLNENYCRNPDGDDAHGPW 420
Db :|||||
Qy 361 YRGNKNTMGNLSQTRSLGTCMDKNMEDLHRHIFWEPDASKLNENYCRNPDGDDAHGPW 420
Db :|||||
Qy 421 CYTGNPLIPWDYCPISRCGEGDTPTIV 447
Db :|||||
Qy 421 CYTGNPLIPWDYCPISRCGEGDTPTIV 447
Db :|||||

RESULT 8
US-07-705-741-1
Sequence 1, Application US/07705741
GENERAL INFORMATION:
APPLICANT: Nakamura, Toshikazu
APPLICANT: Seki, Tatsuya
APPLICANT: Hagiya, Michio
APPLICANT: Shimonishi, Manabu
APPLICANT: Shimizu, Shin
TITLE OF INVENTION: Chromosome DNA Coding For Human
TITLE OF INVENTION: Hepatocyte Growth Factor
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Leydig, Voit & Mayer
STREET: Two Prudential Plaza, Suite 4900
CITY: Chicago
STATE: IL
COUNTRY: U.S.A.
ZIP: 60601-6780
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/705,741
FILING DATE: 19910528
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Green, Robert F.
REGISTRATION NUMBER: 27,555
REFERENCE/DOCKET NUMBER: 38300
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-616-5600
TELEFAX: 312-616-5700
TELEX: 25-3533
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 728 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-07-705-741-1

Query Match 99.9%; Score 2601; DB 3; Length 728;
Best Local Similarity 99.8%; Pred. No. 1e-229;
Matches 446; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ERKRNTTIEFKSAKTLTIKIDPALKIKTKKVNADQCANRCTRNKGLPTCKAFVFDK 60
Db :|||||
Qy 32 QKRRNTTIEFKSAKTLTIKIDPALKIKTKKVNADQCANRCTRNKGLPTCKAFVFDK 91
Db :|||||
Qy 61 ARKQCLWPFPPNSMSGVKKFEGHFDLYENKDYIRNCIIIGKRSYKGTVSTKSGIKQCP 120
Db :|||||
Qy 92 ARKQCLWPFPPNSMSGVKKFEGHFDLYENKDYIRNCIIIGKRSYKGTVSTKSGIKQCP 151
Db :|||||
Qy 121 WSSMIPHEHSLPSSYRGKDLQENYCRNPRGEGPWCFTSNPEVRYEVCIDIPOQSEVEC 180
Db :|||||
Qy 152 WSSMIPHEHSLPSSYRGKDLQENYCRNPRGEGPWCFTSNPEVRYEVCIDIPOQSEVEC 211
Db :|||||
Qy 181 MTCNGESYRGLMDHTESGKICQWRDHTPHRHKFLPERYPDKGDDNYCRNPDGQPRWC 240
Db :|||||
Qy 212 MTCNGESYRGLMDHTESGKICQWRDHTPHRHKFLPERYPDKGDDNYCRNPDGQPRWC 271
Db :|||||
Qy 241 YLDPHTRWECALTKCADNTMNDTVDPLETTECIGOGEGYRGVTNINWGIQCORWDS 300
Db :|||||
Qy 272 YLDPHTRWECALTKCADNTMNDTVDPLETTECIGOGEGYRGVTNINWGIQCORWDS 331
Db :|||||
Qy 301 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 360
Db :|||||
Qy 332 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 391
Db :|||||

```
QY 361 YRGNGKNTMGNSLQSTRSLGTCSMWDMKNNMEDLHRRHIFWEPDASKLNENYCRNPDDDAHGPW 420
Db 392 YRGNGKNTMGNSLQSTRSLGTCSMWDMKNNMEDLHRRHIFWEPDASKLNENYCRNPDDDAHGPW 451
QY 421 CYTGNPLIPWDYCPISRCGEGDTTPTIV 447
Db 452 CYTGNPLIPWDYCPISRCGEGDTTPTIV 478

RESULT 9
US-07-712-932-3
; Sequence 3, Application US/07712932
; GENERAL INFORMATION:
; APPLICANT: Nakamura, Toshikazu
; APPLICANT: Hagiya, Michio
; APPLICANT: Seki, Tatsuya
; APPLICANT: Shimonishi, Manabu
; APPLICANT: Shimizu, Shin
; APPLICANT: Ihara, Izumi
; APPLICANT: Sakaguchi, Mariko
; APPLICANT: Asami, Osamu
; TITLE OF INVENTION: Recombinant Human Hepatocyte Growth
; TITLE OF INVENTION: Factor and Method for Production thereof
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Leydig, Voit & Mayer
; STREET: Two Prudential Plaza, Suite 4900
; CITY: Chicago
; STATE: Illinois
; COUNTRY: U.S.A.
; ZIP: 61801-6780
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/712,932
; FILING DATE: 19910610
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Griffith, Christopher T.
; REGISTRATION NUMBER: 33,392
; REFERENCE/DOCKET NUMBER: 38333
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-616-5600
; TELEFAX: 312-616-5700
; TELEX: 25-3533
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 728 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-712-932-3
Query Match 99.9%; Score 2601; DB 3; Length 728;
Best Local Similarity 99.8%; Pred. No. 1.1e-229;
Matches 446; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ERKRRTTIEFKSAKTTLIKIDPALIKTKKVTADQCANRTRNKGLPFTCKAEVFK 91
Db 32 QRKRRTTIEFKSAKTTLIKIDPALIKTKKVTADQCANRTRNKGLPFTCKAEVFK 91
QY 61 ARKQCLWFPNMSGYKKEGHEFDLYENKDYIRNCLIGKGRSYKGTVSITSGIKCQP 120
Db 92 ARKQCLWFPNMSGYKKEGHEFDLYENKDYIRNCLIGKGRSYKGTVSITSGIKCQP 151
QY 121 WSSMIPHEHSLPSSYRGKOLQENYCRNPRGEGGWCFTSNPEVRYEVCIDIPQCSEVEC 180
Db 152 WSSMIPHEHSLPSSYRGKOLQENYCRNPRGEGGWCFTSNPEVRYEVCIDIPQCSEVEC 211
```

```
QY 181 MTCNGESYRGLMDHTESGKICQRDWDHQTPHRHKFLPERYPDKGFDNDNYCRNPDGQPRPWC 240
Db 212 MTCNGESYRGLMDHTESGKICQRDWDHQTPHRHKFLPERYPDKGFDNDNYCRNPDGQPRPWC 271
QY 241 YTLDPHTRWEYCAIKTCAADNTMNDTDVPLETTECICOGGEGYRGTVNTIWNIGTIPCQRWDS 300
Db 272 YTLDPHTRWEYCAIKTCAADNTMNDTDVPLETTECICOGGEGYRGTVNTIWNIGTIPCQRWDS 331
QY 301 QYPHEHDMTPENFKCKDLRENYCRNPDGSGSPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 360
Db 332 QYPHEHDMTPENFKCKDLRENYCRNPDGSGSPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 391
QY 361 YRGNGKNYMGNSLQSTRSLGTCSMWDMKNNMEDLHRRHIFWEPDASKLNENYCRNPDDDAHGPW 420
Db 392 YRGNGKNYMGNSLQSTRSLGTCSMWDMKNNMEDLHRRHIFWEPDASKLNENYCRNPDDDAHGPW 451
QY 421 CYTGNPLIPWDYCPISRCGEGDTTPTIV 447
Db 452 CYTGNPLIPWDYCPISRCGEGDTTPTIV 478

RESULT 10
US-07-983-366-4
; Sequence 4, Application US/07983366
; GENERAL INFORMATION:
; APPLICANT: SUDO, Tadashi
; APPLICANT: HARADA, Kazumichi
; APPLICANT: HIRAHARA, Ichiro
; APPLICANT: ADACHI, Masami
; TITLE OF INVENTION: NOVEL PROTEIN OF HUMAN ORIGIN AND ITS
; TITLE OF INVENTION: PRODUCTION PROCESS
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Burns, Doane, Swecker & Mathis
; STREET: The George Mason Bldg., Washington & Prince
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/983,366
; FILING DATE: 19921130
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 3-337999
; FILING DATE: 28-NOV-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Crane-Feury, Sharon E
; REGISTRATION NUMBER: 36,113
; REFERENCE/DOCKET NUMBER: 029650-040
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 728 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: YES
US-07-983-366-4
Query Match 99.9%; Score 2601; DB 3; Length 728;
Best Local Similarity 99.8%; Pred. No. 1.1e-229;
Matches 446; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

Qy 1 ERKRNTTHERFKSAKTTLIKIDPALKIKTKKVTADOCANRCKTRNKGLPFTCKAFVDEK 60
Db :|||||
Qy 32 QKRRNTTHERFKSAKTTLIKIDPALKIKTKKVTADOCANRCKTRNKGLPFTCKAFVDEK 91
Db :|||||
Qy 61 ARKQCLWPFNPSMSSGVKKKEGHEFDLYENKDYIRNCIIIGKRSYKGTVSITKSGIKQCP 120
Db :|||||
Qy 92 ARKQCLWPFNPSMSSGVKKKEGHEFDLYENKDYIRNCIIIGKRSYKGTVSITKSGIKQCP 151
Db :|||||
Qy 121 WSSMIPHEHSELPSSYRGKDLQENYCRNPRGEGGPGWCFSTNPEVRYEVCDDIPOCSEVEC 180
Db :|||||
Qy 152 WSSMIPHEHSELPSSYRGKDLQENYCRNPRGEGGPGWCFSTNPEVRYEVCDDIPOCSEVEC 211
Db :|||||
Qy 181 MTCNGESYRGLMDHTESGKICQWRDHQTPHRRKFLPERYPDKGFDNNYCRNPDGQPRPWC 240
Db :|||||
Qy 212 MTCNGESYRGLMDHTESGKICQWRDHQTPHRRKFLPERYPDKGFDNNYCRNPDGQPRPWC 271
Db :|||||
Qy 241 YLDPHTRWYCAIKTCADNTMNDTVPLETTTCIOGOGEGYRGTVNTIWNIGIPCORWDS 300
Db :|||||
Qy 272 YLDPHTRWYCAIKTCADNTMNDTVPLETTTCIOGOGEGYRGTVNTIWNIGIPCORWDS 331
Db :|||||
Qy 301 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 360
Db :|||||
Qy 332 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 391
Db :|||||
Qy 361 YRGNGKNTMGNLSQTRSGLTCSMDKNMEDLHRHIFWEPDASKLNENYCRNPDGQPRPWC 420
Db :|||||
Qy 392 YRGNGKNTMGNLSQTRSGLTCSMDKNMEDLHRHIFWEPDASKLNENYCRNPDGQPRPWC 451
Db :|||||
Qy 421 CYTGNPLIPWDYCPISRCGEGDTTPTIV 447
Db :|||||
Qy 452 CYTGNPLIPWDYCPISRCGEGDTTPTIV 478
Db :|||||

RESULT 11

US-08-242-802-2
; Sequence 2, Application US/08242802
; GENERAL INFORMATION:
; APPLICANT: Nakamura, Toshikazu
; TITLE OF INVENTION: Therapeutic Agent for Renal Disease
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Leydig, Voit & Mayer
; STREET: Two Prudential Plaza, Suite 4900
; CITY: Chicago
; STATE: Illinois
; COUNTRY: U.S.A.
; ZIP: 61801-6780
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/242,802
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/967,562
; FILING DATE:
; APPLICATION NUMBER: US 07/716,792
; FILING DATE: 18-JUN-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Kilyk Jr., John
; REGISTRATION NUMBER: 30,763
; REFERENCE/DOCKET NUMBER: 38351
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-616-5600
; TELEFAX: 312-616-5700
; TELEX: 25-3533
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 728 amino acids
; TYPE: amino acid

; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-242-802-2

Query Match 99.9%; Score 2601; DB 6; Length 728;
Best Local Similarity 99.8%; Pred. No. 1.le-229;
Matches 446; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ERKRNTTHERFKSAKTTLIKIDPALKIKTKKVTADOCANRCKTRNKGLPFTCKAFVDEK 60
Db :|||||
Qy 32 QKRRNTTHERFKSAKTTLIKIDPALKIKTKKVTADOCANRCKTRNKGLPFTCKAFVDEK 91
Db :|||||
Qy 61 ARKQCLWPFNPSMSSGVKKKEGHEFDLYENKDYIRNCIIIGKRSYKGTVSITKSGIKQCP 120
Db :|||||
Qy 92 ARKQCLWPFNPSMSSGVKKKEGHEFDLYENKDYIRNCIIIGKRSYKGTVSITKSGIKQCP 151
Db :|||||
Qy 121 WSSMIPHEHSELPSSYRGKDLQENYCRNPRGEGGPGWCFSTNPEVRYEVCDDIPOCSEVEC 180
Db :|||||
Qy 152 WSSMIPHEHSELPSSYRGKDLQENYCRNPRGEGGPGWCFSTNPEVRYEVCDDIPOCSEVEC 211
Db :|||||
Qy 181 MTCNGESYRGLMDHTESGKICQWRDHQTPHRRKFLPERYPDKGFDNNYCRNPDGQPRPWC 240
Db :|||||
Qy 212 MTCNGESYRGLMDHTESGKICQWRDHQTPHRRKFLPERYPDKGFDNNYCRNPDGQPRPWC 271
Db :|||||
Qy 241 YLDPHTRWYCAIKTCADNTMNDTVPLETTTCIOGOGEGYRGTVNTIWNIGIPCORWDS 300
Db :|||||
Qy 272 YLDPHTRWYCAIKTCADNTMNDTVPLETTTCIOGOGEGYRGTVNTIWNIGIPCORWDS 331
Db :|||||
Qy 301 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 360
Db :|||||
Qy 332 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 391
Db :|||||
Qy 361 YRGNGKNTMGNLSQTRSGLTCSMDKNMEDLHRHIFWEPDASKLNENYCRNPDGQPRPWC 420
Db :|||||
Qy 392 YRGNGKNTMGNLSQTRSGLTCSMDKNMEDLHRHIFWEPDASKLNENYCRNPDGQPRPWC 451
Db :|||||
Qy 421 CYTGNPLIPWDYCPISRCGEGDTTPTIV 447
Db :|||||
Qy 452 CYTGNPLIPWDYCPISRCGEGDTTPTIV 478
Db :|||||

RESULT 12

US-08-292-160B-1
; Sequence 1, Application US/08292160B
; GENERAL INFORMATION:
; APPLICANT: Nakamura, Toshikazu
; APPLICANT: Seki, Tatsuya
; APPLICANT: Hagiya, Michio
; APPLICANT: Shimonishi, Manabu
; APPLICANT: Shimizu, Shin
; TITLE OF INVENTION: Chromosome DNA Coding For Human
; TITLE OF INVENTION: Hepatocyte Growth Factor
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Leydig, Voit & Mayer
; STREET: Two Prudential Plaza, Suite 4900
; CITY: Chicago
; STATE: IL
; COUNTRY: U.S.A.
; ZIP: 60601-6780
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/292,160B
; FILING DATE: 17-AUG-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Green, Robert F.
; REGISTRATION NUMBER: 27,555
; REFERENCE/DOCKET NUMBER: 62748

TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-616-5600
TELEFAX: 312-616-5700
TELEX: 25-3533
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 728 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-292-160B-1

Query Match 99.9%; Score 2601; DB 6; Length 728;
Best Local Similarity 99.8%; Pred. No. 1.1e-229;
Matches 446; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ERKRNTIHEFKSAKTLTIKIDPALKIKTKVNTADQCANRCKTRNKGLPFTCKAFVFDK 60
DB 32 QKRRNTIHEFKSAKTLTIKIDPALKIKTKVNTADQCANRCKTRNKGLPFTCKAFVFDK 91
QY 61 ARKQCLWFPFNSMSSGVKKEFGHEFDLYENKDYIRNCIIIGKRSYKGTYSITKSGIKCOP 120
DB 92 ARKQCLWFPFNSMSSGVKKEFGHEFDLYENKDYIRNCIIIGKRSYKGTYSITKSGIKCOP 151
QY 121 WSSMIPHEHSFLPSSYRGKDLQENYCRNPRGEGGWCFTSNPEVRYEYCDIPQCEVEEC 180
DB 152 WSSMIPHEHSFLPSSYRGKDLQENYCRNPRGEGGWCFTSNPEVRYEYCDIPQCEVEEC 211
QY 181 MTCNGESYRGKDLQENYCRNPRGEGGWCFTSNPEVRYEYCDIPQCEVEEC 240
DB 212 MTCNGESYRGKDLQENYCRNPRGEGGWCFTSNPEVRYEYCDIPQCEVEEC 271
QY 241 YTLDPHTRWEYCAIKTCADNTMDTVPLETTECIGOGEGYRGTVNTIWNIGPCQRWDS 300
DB 272 YTLDPHTRWEYCAIKTCADNTMDTVPLETTECIGOGEGYRGTVNTIWNIGPCQRWDS 331
QY 301 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 360
DB 332 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 391
QY 361 YRGNGKNYMGLSOTRSLGTCMMDKNMEDLHRHIFWEPDASKLNYCRNPDHAGHPW 420
DB 392 YRGNGKNYMGLSOTRSLGTCMMDKNMEDLHRHIFWEPDASKLNYCRNPDHAGHPW 451
QY 421 CYTGNPLIPWDYCPISRCGDTTPTIV 447
DB 452 CYTGNPLIPWDYCPISRCGDTTPTIV 478

RESULT 13
US-08-605-221-2
Sequence 2, Application US/08605221
GENERAL INFORMATION:
APPLICANT: NAKAMURA, Toshikazu
TITLE OF INVENTION: THERAPEUTIC AGENT FOR DISORDER IN BRAIN AND NERVE
FILE REFERENCE: 2520-0101P
CURRENT APPLICATION NUMBER: US/08/605,221
CURRENT FILING DATE: 1996-03-15
NUMBER OF SEQ ID NOS: 4
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 728
TYPE: PRT
ORGANISM: Homo sapiens
OTHER INFORMATION: HGF-A (Full length HGF)
US-08-605-221-2

Query Match 99.9%; Score 2601; DB 10; Length 728;
Best Local Similarity 99.8%; Pred. No. 1.1e-229;
Matches 446; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 ERKRNTIHEFKSAKTLTIKIDPALKIKTKVNTADQCANRCKTRNKGLPFTCKAFVFDK 60

DB 32 QKRRNTIHEFKSAKTLTIKIDPALKIKTKVNTADQCANRCKTRNKGLPFTCKAFVFDK 91
QY 61 ARKQCLWFPFNSMSSGVKKEFGHEFDLYENKDYIRNCIIIGKRSYKGTYSITKSGIKCOP 120
DB 92 ARKQCLWFPFNSMSSGVKKEFGHEFDLYENKDYIRNCIIIGKRSYKGTYSITKSGIKCOP 151
QY 121 WSSMIPHEHSFLPSSYRGKDLQENYCRNPRGEGGWCFTSNPEVRYEYCDIPQCEVEEC 180
DB 152 WSSMIPHEHSFLPSSYRGKDLQENYCRNPRGEGGWCFTSNPEVRYEYCDIPQCEVEEC 211
QY 181 MTCNGESYRGKDLQENYCRNPRGEGGWCFTSNPEVRYEYCDIPQCEVEEC 240
DB 212 MTCNGESYRGKDLQENYCRNPRGEGGWCFTSNPEVRYEYCDIPQCEVEEC 271
QY 241 YTLDPHTRWEYCAIKTCADNTMDTVPLETTECIGOGEGYRGTVNTIWNIGPCQRWDS 300
DB 272 YTLDPHTRWEYCAIKTCADNTMDTVPLETTECIGOGEGYRGTVNTIWNIGPCQRWDS 331
QY 301 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 360
DB 332 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 391
QY 361 YRGNGKNYMGLSOTRSLGTCMMDKNMEDLHRHIFWEPDASKLNYCRNPDHAGHPW 420
DB 392 YRGNGKNYMGLSOTRSLGTCMMDKNMEDLHRHIFWEPDASKLNYCRNPDHAGHPW 451
QY 421 CYTGNPLIPWDYCPISRCGDTTPTIV 447
DB 452 CYTGNPLIPWDYCPISRCGDTTPTIV 478

RESULT 14
US-08-909-515-2
Sequence 2, Application US/08909515
GENERAL INFORMATION:
APPLICANT: KITAMURA, NAOMI
APPLICANT: MIYAZAWA, KEIJI
APPLICANT: DAIKOHARA, YASUSHI
APPLICANT: TSUBOUCHI, HIROHITO
APPLICANT: NAKA, DAIJI
APPLICANT: TAKAHASHI, KAZUHIRO
APPLICANT: MATSUI, RIE
APPLICANT: YOSHIMURA, YOSHIKO
APPLICANT: ISHII, TAKEHISA
TITLE OF INVENTION: HEPATIC PARENCHYMAL CELL GROWTH FACTOR
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN
STREET: 130 WAKE STREET
CITY: BOSTON
STATE: MA
COUNTRY: USA
ZIP: 02019
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/909,515
FILING DATE: 12-AUG-1997
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/455,986
FILING DATE: 31-MAY-1995
APPLICATION NUMBER: 08/089,417
FILING DATE: 09-JUL-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/852,758
FILING DATE: 17-MAR-1992
PRIOR APPLICATION DATA: 07/564,172

;; FILING DATE: 08-AUG-1990
;; ATTORNEY/AGENT INFORMATION:
;; NAME: DAVID, RESNICK S
;; REGISTRATION NUMBER: 34,235
;; REFERENCE/DOCKET NUMBER: 40292-FWC1-DIV
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 617-523-3400
;; TELEFAX: 617-523-6440
;; TELEX: 200291 STRE
;; INFORMATION FOR SEQ ID NO: 2:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 728 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
;; HYPOTHETICAL: NO
;; ANTI-SENSE: NO
;; FRAGMENT TYPE: internal
;; ORIGINAL SOURCE:
;; FEATURE:
US-08-909-515-2

Query Match 99.9%; Score 2601; DB 13; Length 728;
Best Local Similarity 99.8%; Pred. No. 1.le-229;
Matches 446; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ERKRNTIHEFKSAKTLIKIDPALKIKTKKVNADOCANRCKTRNKGLPFTCKAFVFDK 60
Db :|||||
Qy 32 QKRKNTIHEFKSAKTLIKIDPALKIKTKKVNADOCANRCKTRNKGLPFTCKAFVFDK 91
Db :|||||
Qy 61 ARKQCLWFPFNSMSGVKKKEFGHEFDLYENKDYIRNCIIIGKGRSYKGTVSITKSGIKCQP 120
Db :|||||
Qy 92 ARKQCLWFPFNSMSGVKKKEFGHEFDLYENKDYIRNCIIIGKGRSYKGTVSITKSGIKCQP 151
Db :|||||
Qy 121 WSSMIPHEHSLPSSYRGKDQENYCRNPRGEGGPWCFTSNPEVRYEVCIDIPOCSEVEC 180
Db :|||||
Qy 152 WSSMIPHEHSLPSSYRGKDQENYCRNPRGEGGPWCFTSNPEVRYEVCIDIPOCSEVEC 211
Db :|||||
Qy 181 MTCNGESYRGLMDHTESGKICORWDHQTPHRRHKFLPERYPDKGFDNCRNPDGQPRPWC 240
Db :|||||
Qy 212 MTCNGESYRGLMDHTESGKICORWDHQTPHRRHKFLPERYPDKGFDNCRNPDGQPRPWC 271
Db :|||||
Qy 241 YTLDPHTRWYCAIKTCADNTMNDTDVPLETTECIQOGGEGYRGVTNTIWNIPCORWDS 300
Db :|||||
Qy 272 YTLDPHTRWYCAIKTCADNTMNDTDVPLETTECIQOGGEGYRGVTNTIWNIPCORWDS 331
Db :|||||
Qy 301 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 360
Db :|||||
Qy 332 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 391
Db :|||||
Qy 361 YRGNKNTYMGNLSQTRSGLTCSMWDKNMEDLHRHIFWEPDASKLKNENYCRNPDHDAHGPW 420
Db :|||||
Qy 392 YRGNKNTYMGNLSQTRSGLTCSMWDKNMEDLHRHIFWEPDASKLKNENYCRNPDHDAHGPW 451
Db :|||||
Qy 421 CYTGNPLIPWDYCPISRCGGDTTPTIV 447
Db :|||||
Qy 452 CYTGNPLIPWDYCPISRCGGDTTPTIV 478
Db :|||||

RESULT 15
US-08-930-999-1
;; Sequence 1, Application US/08930999
;; GENERAL INFORMATION:
;; APPLICANT: GEMBA, MUNEKAZU
;; APPLICANT: YONEHANA, TSUTOMU
;; TITLE OF INVENTION: AGENT FOR PREVENTING AND/OR TREATING ISCHEMIC DISEASES
;; FILE REFERENCE: 0050-1453-0 PCT
;; CURRENT APPLICATION NUMBER: US/08/930,999
;; CURRENT FILING DATE: 1997-10-21
;; EARLIER APPLICATION NUMBER: PCT/JP96/01065
;; EARLIER FILING DATE: 1996-04-19
;; EARLIER APPLICATION NUMBER: JP 96994/1995

;; EARLIER FILING DATE: 1995-04-21
;; NUMBER OF SEQ ID NOS: 1
;; SOFTWARE: PatentIn Ver. 2.0
;; SEQ ID NO 1
;; LENGTH: 728
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-08-930-999-1

Query Match 99.9%; Score 2601; DB 13; Length 728;
Best Local Similarity 99.8%; Pred. No. 1.le-229;
Matches 446; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ERKRNTIHEFKSAKTLIKIDPALKIKTKKVNADOCANRCKTRNKGLPFTCKAFVFDK 60
Db :|||||
Qy 32 QKRKNTIHEFKSAKTLIKIDPALKIKTKKVNADOCANRCKTRNKGLPFTCKAFVFDK 91
Db :|||||
Qy 61 ARKQCLWFPFNSMSGVKKKEFGHEFDLYENKDYIRNCIIIGKGRSYKGTVSITKSGIKCQP 120
Db :|||||
Qy 92 ARKQCLWFPFNSMSGVKKKEFGHEFDLYENKDYIRNCIIIGKGRSYKGTVSITKSGIKCQP 151
Db :|||||
Qy 121 WSSMIPHEHSLPSSYRGKDQENYCRNPRGEGGPWCFTSNPEVRYEVCIDIPOCSEVEC 180
Db :|||||
Qy 152 WSSMIPHEHSLPSSYRGKDQENYCRNPRGEGGPWCFTSNPEVRYEVCIDIPOCSEVEC 211
Db :|||||
Qy 181 MTCNGESYRGLMDHTESGKICORWDHQTPHRRHKFLPERYPDKGFDNCRNPDGQPRPWC 240
Db :|||||
Qy 212 MTCNGESYRGLMDHTESGKICORWDHQTPHRRHKFLPERYPDKGFDNCRNPDGQPRPWC 271
Db :|||||
Qy 241 YTLDPHTRWYCAIKTCADNTMNDTDVPLETTECIQOGGEGYRGVTNTIWNIPCORWDS 300
Db :|||||
Qy 272 YTLDPHTRWYCAIKTCADNTMNDTDVPLETTECIQOGGEGYRGVTNTIWNIPCORWDS 331
Db :|||||
Qy 301 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 360
Db :|||||
Qy 332 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 391
Db :|||||
Qy 361 YRGNKNTYMGNLSQTRSGLTCSMWDKNMEDLHRHIFWEPDASKLKNENYCRNPDHDAHGPW 420
Db :|||||
Qy 392 YRGNKNTYMGNLSQTRSGLTCSMWDKNMEDLHRHIFWEPDASKLKNENYCRNPDHDAHGPW 451
Db :|||||
Qy 421 CYTGNPLIPWDYCPISRCGGDTTPTIV 447
Db :|||||
Qy 452 CYTGNPLIPWDYCPISRCGGDTTPTIV 478
Db :|||||

Search completed: April 25, 2003, 15:47:54
Job time : 150 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 25, 2003, 15:43:13 ; Search time 36 seconds
(without alignments)
1997.639 Million cell updates/sec

Title: US-09-674-377B-1

Perfect score: 2604

Sequence: 1 ERKRNTHEFKSAKTTLI.....IPWDYCPISRGSGTTPITIV 447

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 731526 seqs, 160883681 residues

Total number of hits satisfying chosen parameters: 731526

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Pending_Patents_AA_New.*

1: /cgn2_6/ptodata/2/paa/PCT_NEW_COMB.pep.*
2: /cgn2_6/ptodata/2/paa/US06_NEW_COMB.pep.*
3: /cgn2_6/ptodata/2/paa/US07_NEW_COMB.pep.*
4: /cgn2_6/ptodata/2/paa/US08_NEW_COMB.pep.*
5: /cgn2_6/ptodata/2/paa/US09_NEW_COMB.pep.*
6: /cgn2_6/ptodata/2/paa/US10_NEW_COMB.pep.*
7: /cgn2_6/ptodata/2/paa/US60_NEW_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	2601	99.9	728	5	US-09-423-516B-2
2	2601	99.9	728	5	US-09-423-516A-2
3	2601	99.9	728	6	US-10-228-206-9
4	2601	99.9	728	7	US-60-453-135-11153
5	2601	99.9	728	7	US-60-453-050-11153
6	2601	99.9	728	7	US-60-455-444-5842
7	2455	94.3	730	6	US-10-311-776-2
8	2417.5	92.8	725	6	US-10-311-776-4
9	1471	56.5	290	6	US-10-283-769-8
10	1180.5	45.3	670	6	US-10-210-130-152
11	1180.5	45.3	711	5	US-09-949-016-6981
12	1180.5	45.3	711	6	US-10-004-378A-118
13	1180.5	45.3	711	6	US-10-004-378A-119
14	1180.5	45.3	711	6	US-10-004-378A-120
15	1168.5	44.9	712	6	US-10-210-130-142
16	1164.5	44.7	543	5	US-09-724-676-63317
17	1164.5	44.7	543	5	US-09-724-676-63328
18	1164.5	44.7	543	5	US-09-724-676A-63317
19	1164.5	44.7	543	5	US-09-724-676A-63328
20	1164.5	44.7	568	5	US-09-724-676-63273
21	1164.5	44.7	568	5	US-09-724-676-63284
22	1164.5	44.7	568	5	US-09-724-676A-63273
23	1164.5	44.7	568	5	US-09-724-676A-63284
24	1164.5	44.7	674	5	US-09-724-676-63251
25	1164.5	44.7	674	5	US-09-724-676-63262
26	1164.5	44.7	674	5	US-09-724-676A-63251

ALIGNMENTS

RESULT 1

US-09-423-516B-2

; Sequence 2, Application US/09423516B

; GENERAL INFORMATION:

; APPLICANT: Medical Research Council

; TITLE OF INVENTION: HGF Polypeptides and their use in therapy

; FILE REFERENCE: 1090-26

; CURRENT APPLICATION NUMBER: US/09/423,516B

; CURRENT FILING DATE: 2000-02-10

; PRIOR APPLICATION NUMBER: PCT/GB98/01318

; PRIOR FILING DATE: 1998-05-07

; PRIOR APPLICATION NUMBER: GB 9709453.6

; PRIOR FILING DATE: 1997-05-10

; NUMBER OF SEQ ID NOS: 3

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 2

; LENGTH: 728

; TYPE: PRT

; ORGANISM: h. sapiens

US-09-423-516B-2

Query Match 99.9%; Score 2601; DB 5; Length 728;

Best Local Similarity 99.8%; Pred. No. 4.4e-206;

Matches 446; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ERKRNTHEFKSAKTTLIKDPAKIKTKKVNADQCANRCKTRNKGLPFTCKAFVFDK 60

DB 32 QKRRTNTHFEKKSATTLIKIDPALKIKTKKVNADQCANRCKTRNKGLPFTCKAFVFDK 91

QY 61 ARKQCLWPFNSMSGVKKEFGHEFDLYENKDYIRNCIIIGKRSYKGTVSITKSGIKCQP 120

DB 92 ARKQCLWPFNSMSGVKKEFGHEFDLYENKDYIRNCIIIGKRSYKGTVSITKSGIKCQP 151

QY 121 WSSMTPHEHSFLPSSYRGKDLQENYCRNPRGEGGPGWCFSTNPEVRYEYCDIPQCSSEVC 180

DB 152 WSSMTPHEHSFLPSSYRGKDLQENYCRNPRGEGGPGWCFSTNPEVRYEYCDIPQCSSEVC 211

QY 181 MTCNSESRYGLMDHYESGKICQWDHQTPIRHKFLPERYPDKGFDNDYCRNPDGQPRPWC 240

DB 212 MTCNSESRYGLMDHYESGKICQWDHQTPIRHKFLPERYPDKGFDNDYCRNPDGQPRPWC 271

QY 241 YTLDPHTREYCAIKTKCADNTMNDTVPLETTTCIQGQEGYRGVTNTWNGIPQCRWDS 300

DB 272 YTLDPHTREYCAIKTKCADNTMNDTVPLETTTCIQGQEGYRGVTNTWNGIPQCRWDS 331

QY 301 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYSQIPLNCDMSHGQDC 360

DB 332 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYSQIPLNCDMSHGQDC 391

Sequence 63262, A
Sequence 63295, A
Sequence 63306, A
Sequence 63295, A
Sequence 63306, A
Sequence 121, App
Sequence 63247, A
Sequence 63246, A
Sequence 63247, A
Sequence 63246, A
Sequence 63247, A
Sequence 63242, A
Sequence 63243, A
Sequence 63243, A
Sequence 63244, A
Sequence 63245, A
Sequence 63244, A
Sequence 63245, A
Sequence 144, App

QY 361 YRGNGKNYGNLSQTRSGLTCSMDKNNMEDLHRRHIFWEPDASKLNENYCRNPDDDAHGWP 420
Db 392 YRGNGKNYGNLSQTRSGLTCSMDKNNMEDLHRRHIFWEPDASKLNENYCRNPDDDAHGWP 451
QY 421 CYTGNPLIPWDYCPISRCGGDTTPTIV 447
Db 452 CYTGNPLIPWDYCPISRCGGDTTPTIV 478

RESULT 2
US-09-423-516A-2
; Sequence 2, Application US/09423516A
; GENERAL INFORMATION:
; APPLICANT: Medical Research Council
; TITLE OF INVENTION: HGF Polypeptides and their use in therapy
; FILE REFERENCE: 1090-26
; CURRENT APPLICATION NUMBER: US/09423,516A
; PRIOR FILING DATE: 2000-02-10
; PRIOR APPLICATION NUMBER: PCT/GB98/01318
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: GB 9709453.6
; PRIOR FILING DATE: 1997-05-10
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 728
; TYPE: PRT
; ORGANISM: h. sapiens
US-09-423-516A-2

Query Match 99.9%; Score 2601; DB 5; Length 728;
Best Local Similarity 99.8%; Pred. No. 4.4e-206;
Matches 446; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ERKRRNTIHEFKSAKTLIKIDPALKIKTKVNTADQCANRTRNKGLPFTCKAFVFDK 60
Db 32 QKRRNTIHEFKSAKTLIKIDPALKIKTKVNTADQCANRTRNKGLPFTCKAFVFDK 91
QY 61 ARKOCLEFPNSMSSGVKKEGHEFDLYENKDYIRNCIIIGKRSYKGTYSITKSGIKCOP 120
Db 92 ARKOCLEFPNSMSSGVKKEGHEFDLYENKDYIRNCIIIGKRSYKGTYSITKSGIKCOP 151
QY 121 WSSMIPHEHSLPSSYRGKDLQENYCRNPRGEGGWCFTSNPEVRYEVCIDIPOCSEVEC 180
Db 152 WSSMIPHEHSLPSSYRGKDLQENYCRNPRGEGGWCFTSNPEVRYEVCIDIPOCSEVEC 211
QY 181 MTCNGESYRGLMDHTESGKICQRDHQTTPHRRKFLPERYDPKGFDDNYCRNPDDGQPRWC 240
Db 212 MTCNGESYRGLMDHTESGKICQRDHQTTPHRRKFLPERYDPKGFDDNYCRNPDDGQPRWC 271
QY 241 YLDPHTRWYCAIKTCADNTMNDTVPLETTICIQGEGYRGTVNTIWNIGIPCORWDS 300
Db 272 YLDPHTRWYCAIKTCADNTMNDTVPLETTICIQGEGYRGTVNTIWNIGIPCORWDS 331
QY 301 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 360
Db 332 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 391
QY 361 YRGNGKNYGNLSQTRSGLTCSMDKNNMEDLHRRHIFWEPDASKLNENYCRNPDDDAHGWP 420
Db 392 YRGNGKNYGNLSQTRSGLTCSMDKNNMEDLHRRHIFWEPDASKLNENYCRNPDDDAHGWP 451
QY 421 CYTGNPLIPWDYCPISRCGGDTTPTIV 447
Db 452 CYTGNPLIPWDYCPISRCGGDTTPTIV 478

RESULT 3
US-10-228-206-9
; Sequence 9, Application US/10228206
; GENERAL INFORMATION:
; APPLICANT: Hitachi LTD.
; TITLE OF INVENTION: HGF-INTERACTING PROTEINS AND USE THEREOF

FILE REFERENCE: H0100932A1
; CURRENT APPLICATION NUMBER: US/10/228,206
; CURRENT FILING DATE: 2002-08-27
; PRIOR APPLICATION NUMBER: US 60/326,423
; PRIOR FILING DATE: 2001-10-03
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 9
; LENGTH: 728
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-228-206-9

Query Match 99.9%; Score 2601; DB 6; Length 728;
Best Local Similarity 99.8%; Pred. No. 4.4e-206;
Matches 446; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ERKRRNTIHEFKSAKTLIKIDPALKIKTKVNTADQCANRTRNKGLPFTCKAFVFDK 60
Db 32 QKRRNTIHEFKSAKTLIKIDPALKIKTKVNTADQCANRTRNKGLPFTCKAFVFDK 91
QY 61 ARKOCLEFPNSMSSGVKKEGHEFDLYENKDYIRNCIIIGKRSYKGTYSITKSGIKCOP 120
Db 92 ARKOCLEFPNSMSSGVKKEGHEFDLYENKDYIRNCIIIGKRSYKGTYSITKSGIKCOP 151
QY 121 WSSMIPHEHSLPSSYRGKDLQENYCRNPRGEGGWCFTSNPEVRYEVCIDIPOCSEVEC 180
Db 152 WSSMIPHEHSLPSSYRGKDLQENYCRNPRGEGGWCFTSNPEVRYEVCIDIPOCSEVEC 211
QY 181 MTCNGESYRGLMDHTESGKICQRDHQTTPHRRKFLPERYDPKGFDDNYCRNPDDGQPRWC 240
Db 212 MTCNGESYRGLMDHTESGKICQRDHQTTPHRRKFLPERYDPKGFDDNYCRNPDDGQPRWC 271
QY 241 YLDPHTRWYCAIKTCADNTMNDTVPLETTICIQGEGYRGTVNTIWNIGIPCORWDS 300
Db 272 YLDPHTRWYCAIKTCADNTMNDTVPLETTICIQGEGYRGTVNTIWNIGIPCORWDS 331
QY 301 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 360
Db 332 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 391
QY 361 YRGNGKNYGNLSQTRSGLTCSMDKNNMEDLHRRHIFWEPDASKLNENYCRNPDDDAHGWP 420
Db 392 YRGNGKNYGNLSQTRSGLTCSMDKNNMEDLHRRHIFWEPDASKLNENYCRNPDDDAHGWP 451
QY 421 CYTGNPLIPWDYCPISRCGGDTTPTIV 447
Db 452 CYTGNPLIPWDYCPISRCGGDTTPTIV 478

RESULT 4
US-60-453-135-11153
; Sequence 11153, Application US/60453135
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: IAKOUBOVA, Olga
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01456
; CURRENT APPLICATION NUMBER: US/60/453,135
; CURRENT FILING DATE: 2003-03-10
; NUMBER OF SEQ ID NOS: 82762
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11153
; LENGTH: 728
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-453-135-11153

Query Match 99.9%; Score 2601; DB 7; Length 728;
Best Local Similarity 99.8%; Pred. No. 4.4e-206;
Matches 446; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy 1 ERKRNTTHEFKSAKTTLIKIDPALKTKTKVNTADOCANCRTRNKGLPFTCKAFVFDK 60
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
32 QRKRNTTHEFKSAKTTLIKIDPALKTKTKVNTADOCANCRTRNKGLPFTCKAFVFDK 91
Qy 61 ARKQCLWFPFNSMSGVKKFEGHFDLYENKDYIRNCLIGKRSYKGTVSTKSGIKQCP 120
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
92 ARKQCLWFPFNSMSGVKKFEGHFDLYENKDYIRNCLIGKRSYKGTVSTKSGIKQCP 151
Qy 121 WSSMIPHEHSLPSSYRGKDLQENYCRNPRGEGGPGWFTSNPEVRYEVCIDIPOCSEVEC 180
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
152 WSSMIPHEHSLPSSYRGKDLQENYCRNPRGEGGPGWFTSNPEVRYEVCIDIPOCSEVEC 211
Qy 181 MTCNGESYRGLMDHTESGKICQWHDQTPHRRKFLPERYPDKGFDNDYCRNPDGQPRWC 240
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
212 MTCNGESYRGLMDHTESGKICQWHDQTPHRRKFLPERYPDKGFDNDYCRNPDGQPRWC 271
Qy 241 YLDPHTRWEYCAIKTCADNTMNDTVPLETECIGOGEGYRGTVNTIWNIGIPCORWDS 300
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
272 YLDPHTRWEYCAIKTCADNTMNDTVPLETECIGOGEGYRGTVNTIWNIGIPCORWDS 331
Qy 301 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSOIPNCDMSHGQDC 360
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
332 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSOIPNCDMSHGQDC 391
Qy 361 YRGNGKNYMGNSLQSTRSLGTCMMDKNMEDLHRHIFWEPDASKLNENYCRNPDDDAHGWP 420
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
392 YRGNGKNYMGNSLQSTRSLGTCMMDKNMEDLHRHIFWEPDASKLNENYCRNPDDDAHGWP 451
Qy 421 CYTGNPLIPWDYCPISRCGEGDTPPTIV 447
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
452 CYTGNPLIPWDYCPISRCGEGDTPPTIV 478
```

```
RESULT 5
US-60-453-050-11153
; Sequence 11153, Application US/60453050
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: LUKE, May
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: CL001457
; CURRENT APPLICATION NUMBER: US/60/453.050
; CURRENT FILING DATE: 2003-03-10
; NUMBER OF SEQ ID NOS: 82762
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11153
; LENGTH: 728
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-453-050-11153
```

```
Query Match 99.9%; Score 2601; DB 7; Length 728;
Best Local Similarity 99.8%; Pred. No. 4.4e-206;
Matches 446; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 1 ERKRNTTHEFKSAKTTLIKIDPALKTKTKVNTADOCANCRTRNKGLPFTCKAFVFDK 60
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
32 QRKRNTTHEFKSAKTTLIKIDPALKTKTKVNTADOCANCRTRNKGLPFTCKAFVFDK 91
Qy 61 ARKQCLWFPFNSMSGVKKFEGHFDLYENKDYIRNCLIGKRSYKGTVSTKSGIKQCP 120
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
92 ARKQCLWFPFNSMSGVKKFEGHFDLYENKDYIRNCLIGKRSYKGTVSTKSGIKQCP 151
Qy 121 WSSMIPHEHSLPSSYRGKDLQENYCRNPRGEGGPGWFTSNPEVRYEVCIDIPOCSEVEC 180
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
152 WSSMIPHEHSLPSSYRGKDLQENYCRNPRGEGGPGWFTSNPEVRYEVCIDIPOCSEVEC 211
Qy 181 MTCNGESYRGLMDHTESGKICQWHDQTPHRRKFLPERYPDKGFDNDYCRNPDGQPRWC 240
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
212 MTCNGESYRGLMDHTESGKICQWHDQTPHRRKFLPERYPDKGFDNDYCRNPDGQPRWC 271
Qy 241 YLDPHTRWEYCAIKTCADNTMNDTVPLETECIGOGEGYRGTVNTIWNIGIPCORWDS 300
```

```
Db 272 YLDPHTRWEYCAIKTCADNTMNDTVPLETECIGOGEGYRGTVNTIWNIGIPCORWDS 331
Qy 301 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSOIPNCDMSHGQDC 360
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
332 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSOIPNCDMSHGQDC 391
Qy 361 YRGNGKNYMGNSLQSTRSLGTCMMDKNMEDLHRHIFWEPDASKLNENYCRNPDDDAHGWP 420
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
392 YRGNGKNYMGNSLQSTRSLGTCMMDKNMEDLHRHIFWEPDASKLNENYCRNPDDDAHGWP 451
Qy 421 CYTGNPLIPWDYCPISRCGEGDTPPTIV 447
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
452 CYTGNPLIPWDYCPISRCGEGDTPPTIV 478
```

```
RESULT 6
US-60-455-444-5842
; Sequence 5842, Application US/60455444
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: BEGOVICH, Ann
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: CL001455
; CURRENT APPLICATION NUMBER: US/60/455.444
; CURRENT FILING DATE: 2003-03-18
; NUMBER OF SEQ ID NOS: 50986
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5842
; LENGTH: 728
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-455-444-5842
```

```
Query Match 99.9%; Score 2601; DB 7; Length 728;
Best Local Similarity 99.8%; Pred. No. 4.4e-206;
Matches 446; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 1 ERKRNTTHEFKSAKTTLIKIDPALKTKTKVNTADOCANCRTRNKGLPFTCKAFVFDK 60
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
32 QRKRNTTHEFKSAKTTLIKIDPALKTKTKVNTADOCANCRTRNKGLPFTCKAFVFDK 91
Qy 61 ARKQCLWFPFNSMSGVKKFEGHFDLYENKDYIRNCLIGKRSYKGTVSTKSGIKQCP 120
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
92 ARKQCLWFPFNSMSGVKKFEGHFDLYENKDYIRNCLIGKRSYKGTVSTKSGIKQCP 151
Qy 121 WSSMIPHEHSLPSSYRGKDLQENYCRNPRGEGGPGWFTSNPEVRYEVCIDIPOCSEVEC 180
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
152 WSSMIPHEHSLPSSYRGKDLQENYCRNPRGEGGPGWFTSNPEVRYEVCIDIPOCSEVEC 211
Qy 181 MTCNGESYRGLMDHTESGKICQWHDQTPHRRKFLPERYPDKGFDNDYCRNPDGQPRWC 240
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
212 MTCNGESYRGLMDHTESGKICQWHDQTPHRRKFLPERYPDKGFDNDYCRNPDGQPRWC 271
Qy 241 YLDPHTRWEYCAIKTCADNTMNDTVPLETECIGOGEGYRGTVNTIWNIGIPCORWDS 300
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
272 YLDPHTRWEYCAIKTCADNTMNDTVPLETECIGOGEGYRGTVNTIWNIGIPCORWDS 331
Qy 301 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSOIPNCDMSHGQDC 360
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
332 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSOIPNCDMSHGQDC 391
Qy 361 YRGNGKNYMGNSLQSTRSLGTCMMDKNMEDLHRHIFWEPDASKLNENYCRNPDDDAHGWP 420
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
392 YRGNGKNYMGNSLQSTRSLGTCMMDKNMEDLHRHIFWEPDASKLNENYCRNPDDDAHGWP 451
Qy 421 CYTGNPLIPWDYCPISRCGEGDTPPTIV 447
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
452 CYTGNPLIPWDYCPISRCGEGDTPPTIV 478
```

RESULT 7

```
US-10-311-776-2
; Sequence 2, Application US/10311776
; GENERAL INFORMATION:
; APPLICANT: MIYAKE, Masashi
; APPLICANT: IWABUCHI, Shigehiro
; APPLICANT: SUZUTA, Yasuyuki
; TITLE OF INVENTION: Canine Hepatocyte Growth Factor
; FILE REFERENCE: 522.1023
; CURRENT APPLICATION NUMBER: US/10/311.776
; PRIOR FILING DATE: 2002-12-18
; PRIOR APPLICATION NUMBER: JP2000-187724
; PRIOR FILING DATE: 2000-06-22
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 2
; LENGTH: 730
; TYPE: PRT
; ORGANISM: Canis familiaris
; FEATURE:
; NAME/KEY: Variation
; LOCATION: (138)
; OTHER INFORMATION: Xaa = Gly
US-10-311-776-2

Query Match          94.3%; Score 2455; DB 6; Length 730;
Best Local Similarity 92.8%; Pred. No. 4.7e-194;
Matches 415; Conservative 19; Mismatches 13; Indels 0; Gaps 0;

QY 1 ERKRRNTIHEFKSAKTLIKIDPALKIKTKVNTADQCANRCTRNKGLPFTCKAFVDEK 60
Db 32 QKRRNTLHEFKSAKTLIKEDPLLLKIKTKMNTADQCANCTIRNKGLPFTCKAFVDEK 91
QY 61 ARKOCLEWFPNSMSSGVKKEGHEFDLYENKDYIRNCIIIGKRSYKGTVSIITKSGIKCOP 120
Db 92 ARKCLWFPNSMTSGYKKEGHEFDLYENKDYIRNCIIIGKGSYKXTVSIITKSGIKCOP 151
QY 121 WSSMIPHEHSLPSSYRGKDLQENYCRNPRGEGGPGWCFTSNPEVRYEYCDIPQCSSEVC 180
Db 152 WNSMIPHEHSLPSSYRGKDLQENYCRNPRGEGGPGWCFTSNPEVRYEYCDIPQCSSEVC 211
QY 181 MTCNGESYRGLMDHTEGSKICQWDHQTPHRHKFLPERYDPKGFDDNYCRNPDGQPRPWC 240
Db 212 MTCNGESYRGLMDHTEGSKICQWDHQTPHRHKFLPERYDPKGFDDNYCRNPDGQPRPWC 271
QY 241 YTLDPHTWEYCAIKTCADNTMDTVPLETTTCIOGGEYRGTVNTIWNIGPCQWRDS 300
Db 272 YTLDPHTWEYCAIKTCADNTMDTVPLETTTCIOGGEYRGTVNTIWNIGPCQWRDS 331
QY 301 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 360
Db 332 QYPHGHDTTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 391
QY 361 YRGNGKNTMGNSLSTRSGLTCSMWKKNMEDLHRIHFWEPPDASKLNENYCRNPDGDAHGWP 420
Db 392 YRGNGKNTMGNSLSTRSGLTCSMWKKNMEDLHRIHFWEPPDASKLNENYCRNPDGDAHGWP 451
QY 421 CYTGNPLIPWDYCPISRCGEGDTTPIV 447
Db 452 CYTGNPLIPWDYCPISRCGEGDTTPIV 478

RESULT 8
US-10-311-776-4
; Sequence 4, Application US/10311776
; GENERAL INFORMATION:
; APPLICANT: MIYAKE, Masashi
; APPLICANT: IWABUCHI, Shigehiro
; APPLICANT: SUZUTA, Yasuyuki
; TITLE OF INVENTION: Canine Hepatocyte Growth Factor
; FILE REFERENCE: 522.1023
; CURRENT APPLICATION NUMBER: US/10/311.776
; CURRENT FILING DATE: 2002-12-18
; PRIOR APPLICATION NUMBER: JP2000-187724

US-10-283-769-8
; Sequence 8, Application US/10283769
; GENERAL INFORMATION:
; APPLICANT: Chan, Andrew ML
; APPLICANT: Rubin, Jeffrey L.
; APPLICANT: Bottaro, Donald P.
; APPLICANT: Aaronson, Stuart A.
; APPLICANT: Stahl, Stephen J.
; APPLICANT: Wingfield, Paul T.
; APPLICANT: Cioce, Vittoria
; TITLE OF INVENTION: TRUNCATED HEPATOCYTE GROWTH FACTOR
; FILE REFERENCE: 14014.0269
; CURRENT APPLICATION NUMBER: US/10/283.769
; CURRENT FILING DATE: 2002-10-30
; PRIOR APPLICATION NUMBER: US/08/484.841
; PRIOR FILING DATE: 2002-03-13
; PRIOR APPLICATION NUMBER: 08/130.134
; PRIOR FILING DATE: 1993-10-04
; PRIOR APPLICATION NUMBER: 07/655.502
; PRIOR FILING DATE: 1991-02-15
; PRIOR APPLICATION NUMBER: 07/582.063

US-10-311-776-4
; PRIOR FILING DATE: 2000-06-22
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 4
; LENGTH: 725
; TYPE: PRT
; ORGANISM: Canis familiaris
; FEATURE:
; NAME/KEY: Variation
; LOCATION: (138)
; OTHER INFORMATION: Xaa = Gly
US-10-311-776-4

Query Match          92.8%; Score 2417.5; DB 6; Length 725;
Best Local Similarity 91.7%; Pred. No. 5.8e-191;
Matches 410; Conservative 19; Mismatches 13; Indels 5; Gaps 1;

QY 1 ERKRRNTIHEFKSAKTLIKIDPALKIKTKVNTADQCANRCTRNKGLPFTCKAFVDEK 60
Db 32 QKRRNTLHEFKSAKTLIKEDPLLLKIKTKMNTADQCANCTIRNKGLEPFTCKAFVDEK 91
QY 61 ARKOCLEWFPNSMSSGVKKEGHEFDLYENKDYIRNCIIIGKRSYKGTVSIITKSGIKCOP 120
Db 92 ARKCLWFPNSMTSGYKKEGHEFDLYENKDYIRNCIIIGKGSYKXTVSIITKSGIKCOP 151
QY 121 WSSMIPHEHSLPSSYRGKDLQENYCRNPRGEGGPGWCFTSNPEVRYEYCDIPQCSSEVC 180
Db 152 WNSMIPHEHSLPSSYRGKDLQENYCRNPRGEGGPGWCFTSNPEVRYEYCDIPQCSSEVC 206
QY 181 MTCNGESYRGLMDHTEGSKICQWDHQTPHRHKFLPERYDPKGFDDNYCRNPDGQPRPWC 240
Db 207 MTCNGESYRGLMDHTEGSKICQWDHQTPHRHKFLPERYDPKGFDDNYCRNPDGQPRPWC 266
QY 241 YTLDPHTWEYCAIKTCADNTMDTVPLETTTCIOGGEYRGTVNTIWNIGPCQWRDS 300
Db 267 YTLDPHTWEYCAIKTCADNTMDTVPLETTTCIOGGEYRGTVNTIWNIGPCQWRDS 326
QY 301 QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 360
Db 327 QYPHGHDTTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 386
QY 361 YRGNGKNTMGNSLSTRSGLTCSMWKKNMEDLHRIHFWEPPDASKLNENYCRNPDGDAHGWP 420
Db 387 YRGNGKNTMGNSLSTRSGLTCSMWKKNMEDLHRIHFWEPPDASKLNENYCRNPDGDAHGWP 446
QY 421 CYTGNPLIPWDYCPISRCGEGDTTPIV 447
Db 447 CYTGNPLIPWDYCPISRCGEGDTTPIV 473

RESULT 9
US-10-283-769-8
; Sequence 8, Application US/10283769
; GENERAL INFORMATION:
; APPLICANT: Chan, Andrew ML
; APPLICANT: Rubin, Jeffrey L.
; APPLICANT: Bottaro, Donald P.
; APPLICANT: Aaronson, Stuart A.
; APPLICANT: Stahl, Stephen J.
; APPLICANT: Wingfield, Paul T.
; APPLICANT: Cioce, Vittoria
; TITLE OF INVENTION: TRUNCATED HEPATOCYTE GROWTH FACTOR
; FILE REFERENCE: 14014.0269
; CURRENT APPLICATION NUMBER: US/10/283.769
; CURRENT FILING DATE: 2002-10-30
; PRIOR APPLICATION NUMBER: US/08/484.841
; PRIOR FILING DATE: 2002-03-13
; PRIOR APPLICATION NUMBER: 08/130.134
; PRIOR FILING DATE: 1993-10-04
; PRIOR APPLICATION NUMBER: 07/655.502
; PRIOR FILING DATE: 1991-02-15
; PRIOR APPLICATION NUMBER: 07/582.063
```

```
; PRIOR FILING DATE: 1990-09-14
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 290
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Artificial Sequence ;/ Note = synthetic construct
US-10-283-769-8

Query Match          56.5%; Score 1471; DB 6; Length 290;
Best Local Similarity 99.6%; Pred. No. 2.3e-113;
Matches 256; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ERKRNTTIEFKKSAKTLIKIDPALKIKTKKVTADQCANCTRNKGLPFTCKAFVFDK 60
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 32 QRKRNTTIEFKKSAKTLIKIDPALKIKTKKVTADQCANCTRNKGLPFTCKAFVFDK 91

QY 61 ARKQCLWFFPNSMSGVKKEGHEFDLYENKDYIRNCIIIGKRSYKGTVITSIKGIKQCP 120
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 92 ARKQCLWFFPNSMSGVKKEGHEFDLYENKDYIRNCIIIGKRSYKGTVITSIKGIKQCP 151

QY 121 WSSMIPHEHSLPSSYRGKDLQENYCRNPRGEGGWCFTSNPEVRYEYCDIPIQCEVEPC 180
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 152 WSSMIPHEHSLPSSYRGKDLQENYCRNPRGEGGWCFTSNPEVRYEYCDIPIQCEVEPC 211

QY 181 MTCNGESYGLMDHTESGKICQWHDQTHPHKFLPERYPDKGFDNNYCRNPDGQPRPWC 240
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 212 MTCNGESYGLMDHTESGKICQWHDQTHPHKFLPERYPDKGFDNNYCRNPDGQPRPWC 271

QY 241 YTLDPHTRWEYCAIKTC 257
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 272 YTLDPHTRWEYCAIKTC 288

RESULT 10
US-10-210-130-152
; Sequence 152, Application US/10210130
; GENERAL INFORMATION:
; APPLICANT: Zerhusen, Bryan D.
; APPLICANT: Patturajan, Meera
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Miller, Charles E.
; APPLICANT: Rieger, Daniel K.
; APPLICANT: Pena, Carol E.A.
; APPLICANT: Shimkets, Richard A.
; APPLICANT: Li, Li
; APPLICANT: Berghs, Constance
; APPLICANT: Zhong, Mei
; APPLICANT: Casman, Stacie J.
; APPLICANT: Voss, Edward Z.
; APPLICANT: Boldog, Ferenc L.
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Smithson, Glennda
; APPLICANT: Ji, Weizhen
; APPLICANT: Gorman, Linda
; APPLICANT: Vernet, Corine A.M.
; APPLICANT: Leite, Mario W.
; APPLICANT: Guo, Xiaojia Sasha
; APPLICANT: Anderson, David W.
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Burgess, Catherine E.
; APPLICANT: Khramtsov, Nikolai V.
; APPLICANT: Ort, Tatiana
; APPLICANT: Ellerman, Karen
; APPLICANT: Rastelli, Luca
; APPLICANT: Agee, Michele L.
; APPLICANT: Chaudhuri, Amitabha
; APPLICANT: Chant, John S.
; APPLICANT: DiPippo, Vincent A.
; APPLICANT: Edinger, Shlomit R.
```

```
; APPLICANT: Eisen, Andrew J.
; APPLICANT: Gangolli, Esha A.
; APPLICANT: Giot, Loic
; APPLICANT: Ooi, Chean Eng
; APPLICANT: Rothenberg, Mark E.
; APPLICANT: Spaderna, Steven K.
; APPLICANT: Hjalt, Tord
; APPLICANT: Liu, Xiaohong
; APPLICANT: Taupier, Raymond J., Jr.
; APPLICANT: Catterton, Elina
; APPLICANT: Shenoy, Suresh G.
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-416C (Cura-716 SMT)
; CURRENT APPLICATION NUMBER: US/10/210,130
; CURRENT FILING DATE: 2002-08-01
; PRIOR APPLICATION NUMBER: 60/309,501
; PRIOR FILING DATE: 2001-08-02
; PRIOR APPLICATION NUMBER: 60/316,508
; PRIOR FILING DATE: 2001-08-31
; PRIOR APPLICATION NUMBER: 60/354,655
; PRIOR FILING DATE: 2002-02-05
; PRIOR APPLICATION NUMBER: 60/310,291
; PRIOR FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: 60/383,887
; PRIOR FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: 60/310,951
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: 60/323,936
; PRIOR FILING DATE: 2001-09-21
; PRIOR APPLICATION NUMBER: 60/381,039
; PRIOR FILING DATE: 2002-05-16
; PRIOR APPLICATION NUMBER: 60/311,292
; PRIOR FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: 60/311,979
; PRIOR FILING DATE: 2001-08-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 369
; SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 152
; LENGTH: 670
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-210-130-152

Query Match          45.3%; Score 1180.5; DB 6; Length 670;
Best Local Similarity 47.4%; Pred. No. 4.6e-89;
Matches 197; Conservative 62; Mismatches 148; Indels 9; Gaps 4;

QY 33 VNTADOCANRCTRNKGLPFTCTKAFVFDKARKOCLMPPFNMSMSGVKKEGHEFDLYENKD 92
   | :|:|:| | | | | :|:|:| | | | | :|:|:| | | | | :|:|:| |
Db 64 VADAEACAGRC----GPLMDCAFRHYNVSSHGCQLLPWTQHSPHTRLRSGRCDLFQKKD 119

QY 93 YIRNCIIIGKRSYKGTVITSIKGIKQCPWSSMTIPHEHSLPSSYRGKDLQENYCRNPRGE 152
   | :|:|:| | | | | :|:|:| | | | | :|:|:| | | | | :|:|:| |
Db 120 YVRTCIMNNGVGYRGTMATTVGGLPCQASHKFPNDKHKTPTLRNG--LEENFCRNPDGD 177

QY 153 EGGPWCFTSNPEVRYEYCDIPIQCEVECTNGESYRGMLMDHTESGKICQWHDQTHPHRH 212
   | :|:|:| | | | | :|:|:| | | | | :|:|:| | | | | :|:|:| |
Db 178 PGGPWCYTTDPAVFCSCGKISCREAACVYMGEEYRGAVDRTESGRCQWDLQHPHQH 237

QY 213 KFLPERYPDKGFDNNYCRNPDGQPRPWCYTLDPHTRWECATKTCADTNMTDTPLETT 272
   | :|:|:| | | | | :|:|:| | | | | :|:|:| | | | | :|:|:| |
Db 238 PFEPGKFLDQGLDDNYCRNPDGSERPWCYTTDPOIEREFCDLPRCSEAQPROEA--TTV 295

QY 273 ECIQOGEGYRGTVNTIWNIGIPCORWDSOYPHEHDMTTPENFKCKDLRENYCRNPDGSESP 332
   | :|:|:| | | | | :|:|:| | | | | :|:|:| | | | | :|:|:| |
Db 296 SCFRGKEGYRGTAHTTAGVPCQWDAQIPHQHRTPEKYACKDLRENYCRNPDGSEAP 355

QY 333 WCFTTDPNIRVGYCSQIPNC--DMSHGQDCYRGNGKNYGNLSQTRSGLTCSMWDKNMEDL 391
   | :|:|:| | | | | :|:|:| | | | | :|:|:| | | | | :|:|:| |
Db 356 WCFTLRPGMRAAFYQIRRCTDDVDPQDCYHGAGEYRGTVSKTRKGVQCORWSAETPHK 415

QY 392 HRHIFWEPDASKLNENYCRNPDODDAHGPWCYTGNNPLIPWDYCPISRCGEGTPTTV 447
```

Db 416 PQFTTSEPHQALEENFCRNPDSHGHCWYCTMDPRTFDFYCALRRACDDQPPSIL 471

RESULT 11

US-09-949-016-6981

Sequence 6981, Application US/09949016

GENERAL INFORMATION:

APPLICANT: VENTER, J. Craig et al.

TITLE OF INVENTION: POLYMERISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

FILE REFERENCE: CLO01307

CURRENT APPLICATION NUMBER: US/09/949,016

CURRENT FILING DATE: 2000-04-14

PRIOR APPLICATION NUMBER: 60/241,755

PRIOR FILING DATE: 2000-10-20

PRIOR APPLICATION NUMBER: 60/237,768

PRIOR FILING DATE: 2000-10-03

PRIOR APPLICATION NUMBER: 60/231,498

PRIOR FILING DATE: 2000-09-08

NUMBER OF SEQ ID NOS: 207012

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 6981

LENGTH: 711

TYPE: PRT

ORGANISM: Human

US-09-949-016-6981

Query Match 45.3%; Score 1180.5; DB 5; Length 711;

Best Local Similarity 47.4%; Pred. No. 4.9e-89;

Matches 197; Conservative 62; Mismatches 148; Indels 9; Gaps 4;

QY 33 VNTADQCANRCTRNKGLPFTCKAFVDFKARKOCLWFPNMSGVKKFEGHFDLYENKD 92

DB 50 VADAEACAGRC-----GPLMDCRAFYHNVSSHGCQQLLPWTQHSPTLRRLRRSGRCDLFQKDD 105

QY 93 YIRNCIIGKRSYKGVSTTKSGIKQCPWSSMIPHEHSLFSPSYRGKDLQENYCRNPGE 152

DB 106 YVRTCIMNNGVGRGTWATTGGLPCQAWSHKFPNDHYKFTPLRNG--LEENFCRNPDDG 163

QY 153 EGGPWCFSTNPEVRYEVCDDIPQCSVEECMTNGESYRGLMDHTESGKICQRDWHTQPHRH 212

DB 164 PGGPWCYTTDPAVRFQSCGKISCREAACWVCNGEYRGAVDRTESGRCQWDLQHPHQH 223

QY 213 KFLPERYPKGGDDNYCRNPDPGQRPWCYTLDPHTRWECYCAIKTCADNTMDTDVPLETT 272

DB 224 PFEFGKFLDQGLDDNYCRNPDSGSRPWCYTTDPTQIEREFCDLPFGCSAQPRQEA--TTV 281

QY 273 ECIQGGEGYRGVTNTIWNIGIPQWRDSDQYPIHEHDMTPENFCKDLRENYCRNPDSGESP 332

DB 282 SCFRGEGYRGVTANTTAGVPCQWRDAQIPHOHRTPEKYACKDLRENFRCRNPDSGSEAP 341

QY 333 WCFTTDPNTRVGCSPINPC-DMSHCQDCYRGNGKYNMGLNLSOTRSGLTCSMWKDNMEDL 391

DB 342 WCFTLRPGMRAECYQIRRCTDDVRQDCQHGAGEYRGVTSKTRGVOCQWRSATPHK 401

QY 392 HRHLFWEPAASKENYCRNPDDAHGWCYTCGNPLIPWDYCPISRCGDDTTPTTV 447

DB 402 PQFTTSEPHQALEENFCRNPDCDSHGHCWYCTMDPRTFDFYCALRRACDDQPPSIL 457

RESULT 12

US-10-004-378A-118

Sequence 118, Application US/10004378A

GENERAL INFORMATION:

APPLICANT: Li, Li

APPLICANT: Furtak, Kazaryna

APPLICANT: Perna, Amanda

APPLICANT: Patturajan, Meera

APPLICANT: Shimkets, Richard A

APPLICANT: Guo, Xiaojia Sasha

APPLICANT: Casman, Stacie J

APPLICANT: Barman, Catherine F

APPLICANT: Malyankar, Uriel M

APPLICANT: Tchernev, Velizar T

APPLICANT: Vernet, Corrine A

APPLICANT: Spytek, Kimberly A

APPLICANT: Agee, Michele

APPLICANT: Rastelli, Luca

APPLICANT: Shenoy, Suresh G

APPLICANT: Grosse, William M

APPLICANT: Alsobrook II, John P

APPLICANT: Lepley, Denise M

APPLICANT: Gerlach, Valerie

APPLICANT: Edinger, Schlomit

APPLICANT: MacDougall, John R

APPLICANT: Peyman, John A

APPLICANT: Gunther, Erik

APPLICANT: Stone, David J

APPLICANT: Ellerman, Karen

APPLICANT: Gangolli, Esha A

TITLE OF INVENTION: Novel Human Proteins, Polynucleotides Encoding Them and Methods of Using the Same

FILE REFERENCE: 21402-179

CURRENT APPLICATION NUMBER: US/10/004,378A

CURRENT FILING DATE: 2001-10-24

PRIOR APPLICATION NUMBER: 60/242,882

PRIOR FILING DATE: 2000-10-24

PRIOR APPLICATION NUMBER: 60/242,765

PRIOR FILING DATE: 2000-10-24

PRIOR APPLICATION NUMBER: 60/300,206

PRIOR FILING DATE: 2001-06-22

PRIOR APPLICATION NUMBER: 60/242,789

PRIOR FILING DATE: 2000-10-24

PRIOR APPLICATION NUMBER: 60/242,768

PRIOR FILING DATE: 2000-10-24

PRIOR APPLICATION NUMBER: 60/242,767

PRIOR FILING DATE: 2000-10-24

PRIOR APPLICATION NUMBER: 60/243,622

PRIOR FILING DATE: 2000-10-26

PRIOR APPLICATION NUMBER: 60/273,047

PRIOR FILING DATE: 2001-03-02

PRIOR APPLICATION NUMBER: 60/243,591

PRIOR FILING DATE: 2000-10-26

PRIOR APPLICATION NUMBER: 60/243,950

PRIOR FILING DATE: 2000-10-27

REMAINING PRIOR APPLICATION DATA REMOVED - SEE FILE WRAPPER OR PALM.

NUMBER OF SEQ ID NOS: 191

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 118

LENGTH: 711

TYPE: PRT

ORGANISM: Homo sapiens

US-10-004-378A-118

Query Match 45.3%; Score 1180.5; DB 6; Length 711;

Best Local Similarity 47.4%; Pred. No. 4.9e-89;

Matches 197; Conservative 62; Mismatches 148; Indels 9; Gaps 4;

QY 33 VNTADQCANRCTRNKGLPFTCKAFVDFKARKOCLWFPNMSGVKKFEGHFDLYENKD 92

DB 50 VADAEACAGRC-----GPLMDCRAFYHNVSSHGCQQLLPWTQHSPTLRRLRRSGRCDLFQKDD 105

QY 93 YIRNCIIGKRSYKGVSTTKSGIKQCPWSSMIPHEHSLFSPSYRGKDLQENYCRNPGE 152

DB 106 YVRTCIMNNGVGRGTWATTGGLPCQAWSHKFPNDHYKFTPLRNG--LEENFCRNPDDG 163

QY 153 EGGPWCFSTNPEVRYEVCDDIPQCSVEECMTNGESYRGLMDHTESGKICQRDWHTQPHRH 212

DB 164 PGGPWCYTTDPAVRFQSCGKISCREAACWVCNGEYRGAVDRTESGRCQWDLQHPHQH 223

QY 213 KFLPERYPKGGDDNYCRNPDPGQRPWCYTLDPHTRWECYCAIKTCADNTMDTDVPLETT 272

DB 224 PFEFGKFLDQGLDDNYCRNPDSGSRPWCYTTDPTQIEREFCDLPFGCSAQPRQEA--TTV 281

QY 273 ECIQGGEGYRGVTNTIWNIGIPQWRDSDQYPIHEHDMTPENFCKDLRENYCRNPDSGESP 332


```

; PRIOR APPLICATION NUMBER: 60/242,765
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: 60/300,206
; PRIOR FILING DATE: 2001-06-22
; PRIOR APPLICATION NUMBER: 60/242,789
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: 60/242,768
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: 60/242,767
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: 60/243,622
; PRIOR FILING DATE: 2000-10-26
; PRIOR APPLICATION NUMBER: 60/273,047
; PRIOR FILING DATE: 2001-03-02
; PRIOR APPLICATION NUMBER: 60/243,591
; PRIOR FILING DATE: 2000-10-26
; PRIOR APPLICATION NUMBER: 60/243,950
; PRIOR FILING DATE: 2000-10-27
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 191
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 120
; LENGTH: 711
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-004-378A-120

Query Match 45.3%; Score 1180.5; DB 6; Length 711;
Best Local Similarity 47.4%; Pred. No. 4.9e-89;
Matches 197; Conservative 62; Mismatches 148; Indels 9; Gaps 4;

Qy 33 VNTADOCANCRTRNKGLEPFTCKAFVFDKARKQCLWFFPNSMSGVKKFEHFDLYENKD 92
Db 50 VADAEACAGRC-----GPLMDCRAFHYNVSSHGCOLLPTQHSPTRLRRSGRCDLFFQKD 105
Qy 93 YIRNCIIIGKRSYKTSITKSGIKCPQWSSMIPHEHSLFSSYRGKDLQENYCRNPRGE 152
Db 106 YVTCIMNNGVYRGTMATTVGGLFCQAWSHKFPNDHYTPTLRNG--LEENFCRNPDGD 163
Qy 153 EGGPWCFSTNPRVYEVCDIPQSEVECMTCNGESYRGMLMBHTESGKICQWDHQTPHRH 212
Db 164 PGGPWCYTTPAVRFQSGIKSCRAACWCNGEYRGAVRTESGRCCQWDLQHPQH 223
Qy 213 KFLPRYDPKGFDDNYCRNPDGQPRWCYTLLDPHTRWYCAIKTCADNTMDTDVPLETT 272
Db 224 PREPGKFLDGLDDNYCRNPDGSRPWCYTTPDQLEREFCDLPCGSEAQRQEA--TTV 281
Qy 273 ECIQGGEGYRGTWNTWNGIPQWRDQSOYPHEHDMTPENFKCKDLRENYCRNPDGSESP 332
Db 282 SCFRKGEGYRGTAHTTAGVPCQWRDAQIPHQHRTPEKYACKDLRENYCRNPDGSEAP 341
Qy 333 WCFTTDPNIRVGYCQIPNC-DMSHGQDCYGRNGKYNMNLSTSGTLTCSMDKNMEDL 391
Db 342 WCFTLRPGMRAAFQIRCTDVRPDQCYHGAGQGYRGTVSKTRKGVQCQWNSAETPHK 401
Qy 392 HRHIFWEPDASKLNENYCRNPDGDAHGWCYTCGNPLIPWDYCPISRCBGTPTTIV 447
Db 402 POFTEFTSEPHQALEENFCRNPDGDSHGWPWCYTMDPRTFDYCALRRCADDPSPSIL 457

RESULT 15
US-10-210-130-142
; Sequence 142, Application US/10210130
; GENERAL INFORMATION:
; APPLICANT: Zerhusen, Bryan D.
; APPLICANT: Patturajan, Meera
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Miller, Charles E.
; APPLICANT: Rieger, Daniel K.
; APPLICANT: Pena, Carol E.A.
; APPLICANT: Shimkets, Richard A.
; APPLICANT: Li, Li
; APPLICANT: Berghs, Constance
```

```

; APPLICANT: Zhong, Mei
; APPLICANT: Casman, Stacie J.
; APPLICANT: Voss, Edward Z.
; APPLICANT: Boldog, Ferenc L.
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Smithson, Glenda
; APPLICANT: Ji, Weizhen
; APPLICANT: Gorman, Linda
; APPLICANT: Vernet, Corine A.M.
; APPLICANT: Leite, Mario W.
; APPLICANT: Guo, Xiaojia Sasha
; APPLICANT: Anderson, David W.
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Burgess, Catherine E.
; APPLICANT: Khramtsov, Nikolai V.
; APPLICANT: Ort, Tatiana
; APPLICANT: Ellerman, Karen
; APPLICANT: Rastelli, Luca
; APPLICANT: Agee, Michele L.
; APPLICANT: Chaudhuri, Amitabha
; APPLICANT: Chant, John S.
; APPLICANT: DiPippo, Vincent A.
; APPLICANT: Edinger, Shlomit R.
; APPLICANT: Eisen, Andrew J.
; APPLICANT: Gangolli, Esha A.
; APPLICANT: Giot, Loic
; APPLICANT: Ooi, Chean Eng
; APPLICANT: Rothenberg, Mark E.
; APPLICANT: Spaderna, Steven K.
; APPLICANT: Hjalt, Tord
; APPLICANT: Liu, Xiaohong
; APPLICANT: Taupier, Raymond J., Jr.
; APPLICANT: Catterton, Elna
; APPLICANT: Shenoy, Suresh G.
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-416C (Cura-716 SMT)
; CURRENT APPLICATION NUMBER: US/10/210,130
; CURRENT FILING DATE: 2002-08-01
; PRIOR APPLICATION NUMBER: 60/309,501
; PRIOR FILING DATE: 2001-08-02
; PRIOR APPLICATION NUMBER: 60/316,508
; PRIOR FILING DATE: 2001-08-31
; PRIOR APPLICATION NUMBER: 60/354,655
; PRIOR FILING DATE: 2002-02-05
; PRIOR APPLICATION NUMBER: 60/310,291
; PRIOR FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: 60/383,887
; PRIOR FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: 60/310,951
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: 60/323,936
; PRIOR FILING DATE: 2001-09-21
; PRIOR APPLICATION NUMBER: 60/381,039
; PRIOR FILING DATE: 2002-05-16
; PRIOR APPLICATION NUMBER: 60/311,292
; PRIOR FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: 60/311,979
; PRIOR FILING DATE: 2001-08-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 369
; SOFTWARE: Curaseqlist version 0.1
; SEQ ID NO 142
; LENGTH: 712
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-210-130-142

Query Match 44.9%; Score 1168.5; DB 6; Length 712;
Best Local Similarity 45.8%; Pred. No. 4.8e-88;
Matches 206; Conservative 73; Mismatches 134; Indels 17; Gaps 8;

Qy 4 RRNTIHEPKSAKTTLIKIDPALKTK--TKKVNADOCANCRTRNKGLEPFTCKAFVFDKA 61
```

Db 19 QRSPLNDFEVLRGTELOQLLOAVVPGMQEDVADAEACAGRC---GPLMDCRAFYHVS 74
Qy 62 RKQCLWFPFNSMGVKKKEFGHEFDLYENKDYIRNCIIIGKGRSYKGTYSITKSGIKCOPW 121
Db 75 SHGCOLLPWTOHSHPRHLRHSGRCDLFOEKDYIRTCINNGVGYRGTMATTVGGLSCQAW 134
Qy 122 SSMIPHEHSFLPSSYRGKDLQENYCRNPRGEGGPMWCFSTNPEVRYEYCDIPQCSEVECM 181
Db 135 SHKEPNDRHYMPTLRNG--LEENFCRNPDPGGPGWCHTTPAVRFQSCGKISCRSAACV 192
Qy 182 TCNGESYGLMDHTESCKICORWDHQTPHRHKFLPERYPDKGFDNDCYCRNPDGQPRWCY 241
Db 193 WCNGEYRGAVDRTESGRECORWDLQHPHQPFBPGKYPDQGLDDNDCYCRNPDGSERPWCY 252
Qy 242 TLDPHTRWEYCAIKTCADNTMNDTVDPLETTECIGQGEYRGVTNTIWNIGIPCORWDSQ 301
Db 253 TTDQIEREFCDLPRGSEAPQROEA--TSVSCFRGKGEYRGVTANTTTAGVPCORWDAQ 310
Qy 302 YPHEHMDTPENFKCKDLRENYCRNPDGSESPWCFSTDNIRVGYCSQIPNC-DMSHGODC 360
Db 311 IPHQHRTPEKYACKDLRENEFCRNPDPGSEAPWCFTLRPGMRVGCYQIRRCTDDVRPOGC 370
Qy 361 YRGCKNYMGNLQSRTSGLTCSMWDKNMEDLHRHFW---EPDASKLNEYCRNPDDDAH 417
Db 371 YHGAGEQYRGTVSKTRKGVCOR--ASAEHPKPOFTTSEPHA-OLEENFCRDPDGSY 427
Qy 418 GPWCYTGNNPLIPWDYCPISRCEGDTTPTIV 447
Db 428 GPWCYTMDPRTPFYCALRCADDOPPSIL 457

Search completed: April 25, 2003, 15:48:37
Job time : 41 secs

THIS PAGE BLANK (uspto)

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 25, 2003, 15:48:43 ; Search time 46 Seconds
(without alignments)
934.176 Million cell updates/sec

Title: US-09-674-377B-1

Perfect score: 2604

Sequence: 1 ERRRNTIHEFKSAKTLI.....IPWDYCPISRCGDTPTIV 447

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_73.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2601	99.9	728	1 JH0579	hepatocyte growth
2	2408	92.5	728	1 A60185	hepatocyte growth
3	2401	92.2	728	1 A35644	hepatocyte growth
4	1864.5	71.6	710	1 I51283	hepatocyte growth
5	1737.5	66.7	411	2 I51285	hepatocyte growth
6	1180.5	45.3	711	1 A47136	macrophage-stimula
7	1144	43.9	716	1 JC5061	macrophage-stimula
8	1125	43.2	716	1 A40332	macrophage-stimula
9	1066	40.9	790	1 PLPG	plasma
10	1059	40.7	812	1 PLMS	plasma
11	1053.5	40.5	810	2 I46260	plasma
12	1044.5	40.1	810	2 B30848	plasma
13	1042	40.0	810	1 PLHU	plasma
14	1038	39.9	812	1 PLBO	plasma
15	943.5	36.2	2869	2 T18518	apolipoprotein(a)
16	879	33.8	4548	1 S00657	apolipoprotein(a)
17	843.5	32.4	1420	2 A32869	apolipoprotein(a)
18	457	17.5	460	2 B61545	plasma
19	454	17.4	455	2 A61545	plasma
20	408.5	15.7	336	2 S33879	plasma
21	357.5	13.7	625	1 TBBO	thrombin
22	353	13.6	169	2 A40522	plasma
23	352.5	13.5	618	2 A35827	thrombin
24	350	13.4	622	1 TBHU	thrombin
25	344	13.2	617	2 S10511	thrombin
26	307.5	11.8	559	1 A35029	t-plasminogen acti
27	303	11.6	562	1 UKHUT	t-plasminogen acti
28	297.5	11.4	559	1 A29941	t-plasminogen acti
29	270	10.4	120	2 B61545	plasma

RESULT 1

JH0579

hepatocyte growth factor precursor [validated] - human

N:Alternate names: hepaoleitin A; scatter factor

C:Species: Homo sapiens (man)

C:Date: 17-Aug-1992 #sequence-revision 17-Aug-1992 #text-change 08-Dec-2000

C:Accession: JH0579; JH0333; A41140; B36677; A35512; A39006; PH0114; A37796;

R:Seki, T.; Hagiya, M.; Shimonishi, M.; Nakamura, T.; Shimizu, S.

Gene 102, 213-219, 1991

A:Title: Organization of the human hepatocyte growth factor-encoding gene.

A:Reference number: JH0579; MUID:91340155; PMID:1831432

A:Accession: JH0579

A:Molecule type: DNA

A:Residues: 1-728 <SEK>

A:Cross-references: DDBJ:D90318

A:Note: The authors translated the codon GAA for residue 662 as Gly

R:Seki, T.; Hagiya, M.; Shimonishi, M.; Nakamura, T.; Shimizu, S.

submitted to JIPID, March 1991

A:Description: Organization of the human hepatocyte growth factor-encoding gene.

A:Reference number: JH0333

A:Accession: JH0333

A:Molecule type: DNA

A:Residues: 1-481, RT', 484-728 <SE2>

R:Weidner, K.M.; Arakaki, N.; Hartmann, G.; Vandekerckhove, J.; Weingart, S.; Rieder,

Proc. Natl. Acad. Sci. U.S.A. 88, 7001-7005, 1991

A:Title: Evidence for the identity of human scatter factor and human hepatocyte growth

A:Reference number: A41140; MUID:91334393; PMID:1831266

A:Accession: A41140

A:Molecule type: mRNA

A:Residues: 1-728 <WEI>

A:Cross-references: GB:M73239; NID:g337935; PIDN:AAA64239.1; PID:g337936

R:Seki, T.; Ihara, I.; Sugimura, A.; Shimonishi, M.; Nishizawa, T.; Asami, O.; Hagiya

Biochem. Biophys. Res. Commun. 172, 321-327, 1990

A:Title: Isolation and expression of cDNA for different forms of hepatocyte growth fa

A:Reference number: A36677; MUID:91025062; PMID:2145836

A:Accession: B36677

A:Molecule type: mRNA

A:Residues: 1-728 <SE3>

A:Cross-references: GB:M60718; NID:g184031; PIDN:AAA52648.1; PID:g184032

A:Accession: A36677

A:Molecule type: mRNA

A:Residues: 1-161, 167-728 <SE4>

A:Cross-references: EMBL:X16323

A:Experimental source: leukocyte

R:Miyazawa, K.; Tsubouchi, H.; Naka, D.; Takahashi, K.; Okigaki, M.; Arakaki, N.; Nak

Biochem. Biophys. Res. Commun. 163, 967-973, 1989

A:Title: Molecular cloning and sequence analysis of cDNA for human hepatocyte growth

A:Reference number: A35512; MUID:89392017; PMID:2528952

A:Accession: A35512

A>Status: not compared with conceptual translation

A:Molecule type: mRNA

A:Residues: 1-728 <MY>

A:Cross-references: GB:M29145; NID:g184041; PIDN:AAA52650.1; PID:g306846
 R:Rubin, J.S.; Chan, A.M.L.; Bottaro, D.P.; Burgess, W.H.; Taylor, W.G.; Cech, A.C.; Hlin
 Proc. Natl. Acad. Sci. U.S.A. 88, 415-419, 1991
 A:Title: A broad-spectrum human lung fibroblast-derived mitogen is a variant of hepatocyte
 A:Reference number: A39006; MUID:91110540; PMID:1824873
 A:Accession: A39006
 A:Molecule type: mRNA
 A:Residues: 1-161,167-728 <RUB>
 A:Cross-references: GB:M53379
 A:Experimental source: embryonic lung
 R:Yoshiyama, Y.; Arahaki, N.; Naka, D.; Takahashi, K.; Hirono, S.; Kondo, J.; Nakayama,
 Biochem. Biophys. Res. Commun. 175, 660-667, 1991
 A:Title: Identification of the N-terminal residue of the heavy chain of both native and
 A:Reference number: PH0114; MUID:91207365; PMID:1826837
 A:Accession: PH0114
 A:Molecule type: protein
 A:Residues: 32-43;53-58 <YOS>
 A:Experimental source: plasma
 R:Weidner, K.M.; Behrens, J.; Vandekerckhove, J.; Birchmeier, W.
 J. Cell Biol. 111, 2097-2108, 1990
 A:Title: Scatter factor: molecular characteristics and effect on the invasiveness of ep
 A:Reference number: A37796; MUID:91035621; PMID:2146276
 A:Accession: A37796
 A:Molecule type: protein
 A:Residues: 86-91;329-344;356-363,'XX',366-370;425-434;442-447,'X',449-450;543-546,'X',5
 R:Nakamura, T.; Nishizawa, T.; Hagiya, M.; Seki, T.; Shimonishi, M.; Sugimura, A.; Tashi
 Nature 342, 440-443, 1989
 A:Title: Molecular cloning and expression of human hepatocyte growth factor.
 A:Reference number: S06794; MUID:90066676; PMID:2531289
 A:Accession: S06794
 A:Molecule type: mRNA
 A:Residues: 1-31,'HK',34-77,'N',79-292,'V',294-299,'M',301-316,'A',318-335,'K',337-386,'
 A:Cross-references: EMBL:X16323; NID:g32081; PIDN:CAA34387.1; PID:g32082
 A:Experimental source: liver
 A:Note: the authors translated the codon CAG for residue 727 as Glu
 A:Note: part of this sequence, including the amino end of both the alpha and beta chains
 R:Hartmann, G.; Naidini, L.; Weidner, K.M.; Sachs, M.; Vigna, E.; Comoglio, P.M.; Birchm
 Proc. Natl. Acad. Sci. U.S.A. 89, 11574-11578, 1992
 A:Title: A functional domain in the heavy chain of scatter factor/hepatocyte growth fact
 A:Reference number: I59214; MUID:93087571; PMID:1280830
 A:Accession: I59214
 A:Status: preliminary; translated from GB/EMBL/DDBJ
 A:Molecule type: mRNA
 A:Residues: 1-288,'ET' <HAR>
 A:Cross-references: GB:L02931; NID:g184033; PIDN:AAA52649.1; PID:g184034
 R:Miyazawa, K.; Kitamura, A.; Naka, D.; Kitamura, N.
 Eur. J. Biochem. 197, 15-22, 1991
 A:Title: An alternatively processed mRNA generated from human hepatocyte growth factor
 A:Reference number: S15443; MUID:91200041; PMID:1826653
 A:Accession: S15443
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-288,'ET' <MIY2>
 A:Cross-references: EMBL:X57574; NID:g32083; PIDN:CAA40802.1; PID:g32084
 R:Shima, N.; Nagao, M.; Ogaki, F.; Tsuda, E.; Murakami, A.; Higashio, K.
 Biochem. Biophys. Res. Commun. 180, 1151-1158, 1991
 A:Title: Tumor cytotoxic factor/hepatocyte growth factor from human fibroblasts: cloning
 A:Reference number: I52253; MUID:92062058; PMID:1835383
 A:Accession: I52253
 A:Status: preliminary; translated from GB/EMBL/DDBJ
 A:Molecule type: mRNA
 A:Residues: 161-166 <SHI>
 A:Cross-references: GB:S62561; NID:g237996; PIDN:AAB20169.1; PID:g237997
 A:Gene: GDB:HGF
 A:Cross-references: GDB:127524; OMIM:142409
 A:Map position: 7q21.1-7q21.1
 A:Introns: 30/1; 85/2; 123/1; 161/2; 209/1; 249/2; 289/1; 347/2; 390/1; 424/2; 469/1; 48
 C:Complex: disulfide-bonded heterodimer of chains derived from the same precursor
 C:Function:
 A:Description: stimulates mitosis of hepatocytes and other cells
 A:Note: does not have proteinase activity
 C:Superfamily: hepatocyte growth factor; kringle homology; trypsin homology

C:Keywords: alternative splicing; glycoprotein; growth factor; heterodimer; kringle;
 F:31/Domain: signal sequence #status predicted <SIG>
 F:32-494,495-728/Product: hepatocyte growth factor #status experimental <MAT>
 F:32-494/Domain: alpha chain #status experimental <ACH>
 F:128-206/Domain: kringle homology <KR1>
 F:211-288/Domain: kringle homology <KR2>
 F:305-383/Domain: kringle homology <KR3>
 F:391-469/Domain: kringle homology <KR4>
 F:495-728/Domain: beta chain #status experimental <BCH>
 F:495-716/Domain: trypsin homology <TRY>
 F:32/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status experi
 F:294,402,566,653/Binding site: carboxydrate (Asn) (covalent) #status predicted
 F:487-604/Disulfide bonds: #status predicted

Query Match 99.9%; Score 2601; DB 1; Length 728;
 Best Local Similarity 99.8%; Pred. No. 2,1e-177;
 Matches 446; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ERKRNTTHEFKSAKTLIKIDPALKIKYKYNVADOCANRCTRNGKLPFTCKAFVFK 60
 :|||||
 Db 32 QKRKRNTTHEFKSAKTLIKIDPALKIKYKYNVADOCANRCTRNGKLPFTCKAFVFK 91
 :|||||
 QY 61 ARKOCLEFPENSSGVKKFEGHFDLYENKDYIRNCIIGKGRSYKGTVSITKSGIKCQP 120
 :|||||
 Db 92 ARKOCLEFPENSSGVKKFEGHFDLYENKDYIRNCIIGKGRSYKGTVSITKSGIKCQP 151
 :|||||
 QY 121 WSSMIPHEHSELPSSYRGKDLQENYCRNPRGEGGPMCFSTNPEVRYEVCDDIPOCSEVEC 180
 :|||||
 Db 152 WSSMIPHEHSELPSSYRGKDLQENYCRNPRGEGGPMCFSTNPEVRYEVCDDIPOCSEVEC 211
 :|||||
 QY 181 MTCNGESYRGLMDHTESKICORWDHQTQPHRKELPERYPDKGPDNNYCRNPDQOPRWC 240
 :|||||
 Db 212 MTCNGESYRGLMDHTESKICORWDHQTQPHRKELPERYPDKGPDNNYCRNPDQOPRWC 271
 :|||||
 QY 241 YLDPHTRWECYCAKTCADNTMNDTVDPLETTECIQGGEGYRGTVNTIWMGIPCQRWDS 300
 :|||||
 Db 272 YLDPHTRWECYCAKTCADNTMNDTVDPLETTECIQGGEGYRGTVNTIWMGIPCQRWDS 331
 :|||||
 QY 301 QYPHEHMTPEKFKDLRENYCRNPDGSESPWCFSTNPEVRYEVCDDIPOCSEVEC 360
 :|||||
 Db 332 QYPHEHMTPEKFKDLRENYCRNPDGSESPWCFSTNPEVRYEVCDDIPOCSEVEC 391
 :|||||
 QY 361 YRGNKYNMGNLQSTRGLTCSMDKNNMEDLHRHIFWEPDASKLNENYCRNPDQOPRWC 420
 :|||||
 Db 392 YRGNKYNMGNLQSTRGLTCSMDKNNMEDLHRHIFWEPDASKLNENYCRNPDQOPRWC 451
 :|||||
 QY 421 CYTGNPLIPWDYCPISRCGGDTTPTIV 447
 :|||||
 Db 452 CYTGNPLIPWDYCPISRCGGDTTPTIV 478

RESULT 2
 A60185
 hepatocyte growth factor precursor - mouse
 N:Alternate names: heparin A; scatter factor
 C:Species: Mus musculus (house mouse)
 C:Date: 03-Mar-1993 #sequence revision 26-May-1994 #text change 16-Jun-2000
 C:Accession: J02117; PC2064; A60185; S43416; S45521; S17173; S10966; I48758; JU0231
 R:Sasaki, M.; Nishio, M.; Sasaki, T.; Enami, J.
 Biochem. Biophys. Res. Commun. 199, 772-779, 1994
 A:Title: Identification of mouse mammary fibroblast-derived mammary growth factor as
 A:Reference number: J02117; MUID:94183257; PMID:8135822
 A:Accession: J02117
 A:Molecule type: mRNA
 A:Residues: 1-728 <SAS2>
 A:Cross-references: GB:D10212; NID:g220435; PIDN:BA01064.1; PID:g220436
 A:Experimental source: fibroblast, COS-1 cell
 A:Note: submitted to JIPID, May 1993
 A:Accession: PC2064
 A:Molecule type: protein
 A:Residues: 496-504 <SAS2>
 R:Rosen, E.M.; Meromsky, L.; Setter, E.; Vinter, D.W.; Goldberg, I.D.
 Proc. Soc. Exp. Biol. Med. 195, 34-43, 1990

Qy	61	ARKOCLMPFPNSMSGVKKDFEGHEFDLYENKDYIRNCIIIGKGRSYKGTVTSITKSGIKCQP	120
Db	93	SRKRKYWYFPNSMSGVKKDFEGHEFDLYENKDYIRNCIIIGKGGYKGTVTSITKSGIKCQP	152
Qy	121	WSSMLPHEHGFPLSSYRGKDLQENYCRNPRGEGGWCFTSNPEVRYEVCDDIPQCSEVEC	180
Db	153	WNSMLPHEHGFPLSSYRGKDLQENYCRNPRGEGGWCFTSNPEVRYEVCDDIPQCSEVEC	212
Qy	181	MTCNGESYRGLMDHTESGKTCQRMWDHQTPIRHKFLPERYPDKGFDNNYCRNPDGQPRPMC	240
Db	213	MTCNGESYRGPMDHTESGKTCQRMWDHQTPIRHKFLPERYPDKGFDNNYCRNPDGKPRPMC	272
Qy	241	YTLDPHTRWEYCAIKTCADNTMNDTVPLETTTCIOGGEGYRGTVNTIWNIGIPCORWDS	300
Db	273	YTLDPDTPWEYCAIKTCAHSANVETDVPMTTTCIOGGEGYRGTVNTIWNIGIPCORWDS	332
Qy	301	QYPHEHDMTPENFCKDKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDDMSHGQDC	360
Db	333	QYPHKHDLITPENFCKDKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPKCDVSSGQDC	392
Qy	361	YRGNKGYMGNLSOTRSGSLTCSMWKKNMEDLHRHIFWEPDASKLNNYCRNPDDDAHGPW	420
Db	393	YRGNKGYMGNLSKTRSGSLTCSMWKKNMEDLHRHIFWEPDASKLNNYCRNPDDDAHGPW	452
Qy	421	CYTGNPLIPWDYCPISRCGEGTTPTV	447
Db	453	CYTGNPLIPWDYCPISRCGEGTTPTV	479
RESULT 3			
A35644			
hepatocyte growth factor precursor - rat			
N:Alternate names: hepaopietin A; scatter factor			
C:Species: Rattus norvegicus (Norway rat)			
C>Date: 28-Sep-1990 #sequence_revision 18-Nov-1992 #text_change 21-Jul-2000			
C:Accession: A35644; S13211			
R:Tashiro, K.; Hagiya, M.; Nishizawa, T.; Seki, T.; Shimonishi, M.; Shimizu			
Proc. Natl. Acad. Sci. U.S.A. 87, 3200-3204, 1990			
A:Title: Deduced primary structure of rat hepatocyte growth factor and exp			
A:Reference number: A35644; MUID:90222197; PMID:2139229			
A:Accession: A35644			
A>Status: preliminary			
A:Molecule type: mRNA			
A:Residues: 1-728 <RAS>			
A:Cross-references: GB:D90102; GB:M32987; NID:9220766; PIDN:BAA14133.1; PI			
A:Note: The authors translated the codon GAC for residue 70 as Gln, GAC fo			
E:Okajima, A.; Miyazawa, K.; Kitamura, N.			
Eur. J. Biochem. 193, 375-381, 1990			
A:Title: Primary structure of rat hepatocyte growth factor and induction o			
A:Reference number: S13211; MUID:91031482; PMID:2146117			
A:Accession: S13211			
A>Status: preliminary			
A:Molecule type: mRNA			
A:Residues: 1-728 <OKA>			
A:Cross-references: EMBL:Y54400; NID:956353; PIDN:CAA38266.1; PID:g4539554			
C:Complex: disulfide-bonded heterodimer of chains derived from the same pr			
C:Function:			
A:Description: stimulates mitosis of hepatocytes and other cells			
A:Note: does not have proteinase activity			
C:Superfamily: hepatocyte growth factor; kringle homology; trypsin homology			
C:Keywords: alternative splicing; glycoprotein; growth factor; heterodimer			
F:1-32/Domain: signal sequence #status predicted <SIG>			
F:56-495/Product: hepatocyte growth factor #status predicted <WAT>			
F:56-495/Domain: hepatocyte growth factor alpha chain #status predicted <A			
F:129-207/Domain: kringle homology <KR1>			
F:122-289/Domain: kringle homology <KR2>			
F:306-384/Domain: kringle homology <KR3>			
F:392-470/Domain: kringle homology <KR4>			
F:496-728/Domain: hepatocyte growth factor			
F:496-719/Domain: trypsin homology <TRY>			
F:33/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #sta			
F:295,403,569,656/Binding site: carboxylate (Asn) (covalent) #status pred			
F:488-607/Disulfide bonds: #status predicted			

```
Query Match          92.2%; Score 2401; DB 1; Length 728;
Best Local Similarity 90.4%; Pred. No. 3.5e-163;
Matches 404; Conservative 23; Mismatches 20; Indels 0; Gaps 0;

QY 1 ERKRNTIHEFKSAKTLIKIDPALKIKTKYNTADQCANRCTRNGKGLPFTCKAFVFDK 60
DB 33 OKKRNTIHEFKSAKTLIKIDPALKIKTKYNTADQCANRCTRNGKGLPFTCKAFVFDK 92
QY 61 ARKQCLWFFPNSMSSGVKKFEGHEFDLYENKDYIRNCIIIGKRSYKGTVITKSGIKCQP 120
DB 93 SRKRCYWPNSMSSGVKKFEGHEFDLYENKDYIRNCIIIGKRSYKGTVITKSGIKCQP 152
QY 121 WSMIPHEHSLPSSYRGKDLQENYCRNPGEGGPGWCFSTNPEVRYEVCIDIQCSEVEC 180
DB 153 WNSMIPHEHSLPSSYRGKDLQENYCRNPGEGGPGWCFSTNPEVRYEVCIDIQCSEVEC 212
QY 181 MTCNGESYRGLMDHTESGKICQRWDHOTPHRHKFLPERYPDKGFDNNYCRNPDGQPRWC 240
DB 213 MTCNGESYRGLMDHTESGKICQRWDHOTPHRHKFLPERYPDKGFDNNYCRNPDGQPRWC 272
QY 241 YLDPHTRWYCAIKTCADNTMDTVPLETTTCIQGGEGYRGTVNTIWNIGIPQCRWDS 300
DB 273 YLDPHTRWYCAIKTCADNTMDTVPLETTTCIQGGEGYRGTVNTIWNIGIPQCRWDS 332
QY 301 QYPHEHMTPEKFKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGDC 360
DB 333 QYPHEHMTPEKFKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGDC 392
QY 361 YRGNKNTYMGNSLQTRSGLTCSMDKNNMEDLHRHIFWEPDASKLNENYCRNPDGDAIGPW 420
DB 393 YRGNKNTYMGNSLQTRSGLTCSMDKNNMEDLHRHIFWEPDASKLNENYCRNPDGDAIGPW 452
QY 421 CYTGNPLIPWDYCPISRCGGDTTPIV 447
DB 453 CYTGNPLIPWDYCPISRCGGDTTPIV 479

RESULT 4
I51283
hepatocyte growth factor precursor - clawed frog
N:Alternate names: hepatopoietin A; scatter factor
C:Species: Xenopus sp. (clawed frog)
C>Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 18-Jun-1999
C:Accession: I51283
R:Nakamura, H.; Tashiro, K.; Nakamura, T.; Shiokawa, K.
Mech. Dev. 49, 123-131, 1995
A:Title: Molecular cloning of Xenopus HGF cDNA and its expression studies in Xenopus ear
A:Reference number: I51283; MUID:95267690; PMID:7748783
A:Accession: I51283
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-710 <N>
A:Cross-references: GB:S77422; NID:9989932; PIDN:AB34354.1; PID:g998933
A:Note: the authors' translation for residue 458 (Thr) is inconsistent with the nucleotide
C:Complex: disulfide-bonded heterodimer of chains derived from the same precursor
C:Function:
A:Description: stimulates mitosis of hepatocytes and other cells
A:Note: does not have proteinase activity
C:Superfamily: hepatocyte growth factor; kringle homology; trypsin homology
F:42-477/478-709/product: hepatocyte growth factor #status predicted <N>
F:42-477/Domain: hepatocyte growth factor alpha chain #status predicted <ACH>
F:115-193/Domain: kringle homology <KRI>
F:198-275/Domain: kringle homology <KR2>
F:289-367/Domain: kringle homology <KR3>
F:375-453/Domain: kringle homology <KR4>
F:478-709/Domain: hepatocyte growth factor beta chain #status predicted <BCH>
F:478-700/Domain: trypsin homology <TRY>
F:52,128,281,322,379,550,637,666/Binding site: carbohydrate (Asn) (covalent) #status pre
F:470-588/Disulfide bonds: #status predicted

Query Match          71.6%; Score 1864.5; DB 1; Length 710;
```

```
Best Local Similarity 69.3%; Pred. No. 4.8e-125;
Matches 305; Conservative 64; Mismatches 68; Indels 3; Gaps 1;

QY 3 KRNTIHEFKSAKTLIKIDPALKIKTKYNTADQCANRCTRNGKGLPFTCKAFVFDKAR 62
DB 21 KKRNAFDYKTAETTLRLNKALEVTKTMFMFTTNCARCSRNGKGLPFTCKAFADKNI 80
QY 63 KQCLWFFPNSMSSGVKKFEGHEFDLYENKDYIRNCIIIGKRSYKGTVITKSGIKCQPS 122
DB 81 KRCHWFSFNTMSAGIKDYIDISFDLYEKDYIRDCIHGKGSNYRTRNVNTRKGLACQPN 140
QY 123 SMIPHEHSLPSSYRGKDLQENYCRNPGEGGPGWCFSTNPEVRYEVCIDIQCSEVECMT 182
DB 141 SMIPHEHSLPSSYRGKDLQENYCRNPGEGGPGWCFSTNPEVRYEVCIDIQCSEVECMT 200
QY 183 CNGESYRGLMDHTESGKICQRWDHOTPHRHKFLPERYPDKGFDNNYCRNPDGQPRWCYT 242
DB 201 CNGESYRGLMDHTESGKICQRWDHOTPHRHKFLPERYPDKGFDNNYCRNPDGQPRWCYT 260
QY 243 LDPHTRWYCAIKTCADNTMDTVPLETTTCIQGGEGYRGTVNTIWNIGIPQCRWDSOY 302
DB 261 LDPHTRWYCAIKTCADNTMDTVPLETTTCIQGGEGYRGTVNTIWNIGIPQCRWDSOY 317
QY 303 PHEHMTPEKFKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGDCYR 362
DB 318 PHLHNTPEKFKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGDCYR 377
QY 363 GNGKNTYMGNSLQTRSGLTCSMDKNNMEDLHRHIFWEPDASKLNENYCRNPDGDAIGPW 422
DB 378 GNGKNTYMGNSLQTRSGLTCSMDKNNMEDLHRHIFWEPDASKLNENYCRNPDGDAIGPW 437
QY 423 TGNPLIPWDYCPISRCGGDT 442
DB 438 TDDPFPWDYCPISRCGGDT 457

RESULT 5
I51285
hepatocyte growth factor/scatter factor - chicken (fragment)
C:Species: Gallus gallus (chicken)
C>Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 16-Jul-1999
C:Accession: I51285
R:Streit, A.; Stern, C.D.; Thery, C.; Ireland, G.W.; Aparicio, S.; Sharpe, M.J.; Gher
Development 121, 813-824, 1995
A:Title: A role for HGF/SF in neural induction and its expression in Hensen's node du
A:Reference number: I51285; MUID:95237013; PMID:7720585
A:Accession: I51285
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-411 <STR>
A:Cross-references: GB:S77480; NID:g998675; PID:g998676
C:Superfamily: hepatocyte growth factor; kringle homology; trypsin homology
F:124-197/Domain: kringle homology <KR>
F:202-279/Domain: kringle homology <KR2>
F:296-374/Domain: kringle homology <KR3>

Query Match          66.7%; Score 1737.5; DB 2; Length 411;
Best Local Similarity 74.9%; Pred. No. 2.9e-116;
Matches 289; Conservative 47; Mismatches 45; Indels 5; Gaps 1;

QY 3 KRNTIHEFKSAKTLIKIDPALKIKTKYNTADQCANRCTRNGKGLPFTCKAFVFDKAR 62
DB 30 KRRNPLHDYKKTGELMLIKVNTLEVKTLNLTTEQCAKRCRNKGLSFTCKAFAYDRVT 89
QY 63 KQCLWFFPNSMSSGVKKFEGHEFDLYENKDYIRNCIIIGKRSYKGTVITKSGIKCQPS 122
DB 90 KRCHWLSFNTMSAGIKDYIDISFDLYEKDYIRDCIHGKGSNYRTRNVNTRKGLACQPN 149
QY 123 SMIPHEHSLPSSYRGKDLQENYCRNPGEGGPGWCFSTNPEVRYEVCIDIQCSEVECMT 182
DB 150 SMIPHEHSLPSSYRGKDLQENYCRNPGEGGPGWCFSTNPEVRYEVCIDIQCSEVECMT 204
QY 183 CNGESYRGLMDHTESGKICQRWDHOTPHRHKFLPERYPDKGFDNNYCRNPDGQPRWCYT 242
```

```

Db 205 CNGESYRGPMDDHTSGECORWDLQRPKHKFRPERYPDKGFDNYCRNPDGKLRPMCYT 264
QY 243 LDPTRWEYCAIKTCADNTMDTVDPLETTCIOGEGYRGVTNTIWNIPCORWDSOY 302
Db 265 LDPNTPWEFCAIKTCVGLNSTEAVETTCIOGEGYRGVTNTIWSGLQCORWDSQF 324
QY 303 PHEHDMTPENFKDLRENYCRNPDGSESPWCFCTTDNIRNIRVGYCSQIPNCDMSHGQDCYR 362
Db 325 PHOHNITPENFKDLRENYCRNPDGSESPWCFCTTDNIRNIRVGYCSQIPKCDVSNQDCYR 384
QY 363 GNGKYNWGNLSQTRSGLTCSMDKNNM 388
Db 385 GNGKSYMGNLSKTRFGLTCSMDKNNM 410

RESULT 6
A47136
macrophage-stimulating protein 1 precursor - human
C:Species: Homo sapiens (man)
C:Date: 03-May-1994 #sequence_revision 14-Nov-1997 #text_change 18-Jun-1999
C:Accession: A40331; A47136; A61395
R:Han, S.; Stuart, L.A.; Degen, S.J.F.
Biochemistry 30, 9768-9780, 1991
A:Title: Characterization of the DNF15S2 locus on human chromosome 3: identification of
A:Reference number: A40331; MUID:92002016; PMID:1655021
A:Accession: A40331
A:Molecule type: DNA
A:Residues: 1-711 <HA1>
A:Cross-references: GB:M74179
A:Accession: B40331
A:Molecule type: mRNA
A:Residues: 1-711 <HA2>
A:Cross-references: GB:M74178; NID:g183976; PIDN:AAA50165.1; PID:g183977
R:Yoshimura, T.; Yuhki, N.; Wang, M.H.; Skeel, A.; Leonard, E.J.
J. Biol. Chem. 268, 15461-15468, 1993
A:Title: Cloning, sequencing, and expression of human macrophage stimulating protein (MS
A:Reference number: A47136; MUID:93340141; PMID:8393443
A:Accession: A47136
A:Molecule type: mRNA
A:Residues: 1-12,'C',14-622,'F',624-711 <YOS>
A:Cross-references: GB:L11924; NID:g398037; PIDN:AAA59872.1; PID:g398038
A:Note: authors translated the codon TTT for residue 623 as Leu; parts of this sequence
R:Skeel, A.; Yoshimura, T.; Showalter, S.B.; Tanaka, S.; Appella, E.; Leonard, E.J.
J. Exp. Med. 173, 1227-1234, 1991
A:Title: Macrophage stimulating protein: purification, partial amino acid sequence, and
A:Reference number: A61395; MUID:91217635; PMID:1827141
A:Accession: A61395
A:Molecule type: protein
A:Residues: 230-247;288-291,'E',293-295,'X',297-301,'X',303,'E',305,'EX',308-310;326-331
A:Experimental source: plasma
C:Genetics:
A:Gene: GDB:MST1; D3F15S2; DNF15S2; HGFL
A:Cross-references: GDB:128833; OMIM:142408
A:Map position: 3p21-3p21.3
C:Complex: disulfide-bonded heterodimer of chains derived from the same precursor
C:Superfamily: hepatocyte growth factor; kringle homology; trypsin homology
C:Keywords: duplication; glycoprotein; growth factor; kringle; plasma
F:1-18/Domain: signal sequence #status predicted <SIG>
F:19-483,484-711/Product: macrophage-stimulating protein 1 #status predicted <MAT>
F:19-483/Domain: alpha chain #status predicted <ACH>
F:110-186/Domain: kringle homology <KR1>
F:191-268/Domain: kringle homology <KR2>
F:283-361/Domain: kringle homology <KR3>
F:370-448/Domain: kringle homology <KR4>
F:484-711/Domain: beta chain #status predicted <BCH>
F:484-704/Domain: trypsin homology <TRY>
F:56-78,60-66;110-186,131-169,157-181,191-268,212-251,240-263,283-361,304-343,332-355,37
F:72,296,615/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 45.38; Score 1180.5; DB 1; Length 711;
Best Local Similarity 47.48; Pred. No. 2.1e-76;
Matches 197; Conservative 62; Mismatches 148; Indels 9; Gaps 4;

```

```

QY 33 VNTADQCANRCTRNRKGLPFTCKAFVDFKARKQCLWFFPNSMSSGVYKKEFGHEFDLYENKD 92
Db 50 VADAEACAGRC---GPLMDCRAFHYNVSSHGCOLLPWTQHSPTLRLRRSGRCDLFQKKD 105
QY 93 YIRNCIICKGRSYKGTIVTSIKSGIKCQPWSSMIPEHSFLPSSYRGKDLQENYCRNPRGE 152
Db 106 YVRTCIMNNGVGYRTMATTVGGPLCQAWSHKFFNDHKYTTPLRNG--LEENFCRNPDGD 163
QY 153 EGGPWCFSTNPEVRYEVCIDIPOCSEVCMTCNGESYRGKGLMDHTESGKICQRDWDHQTPIRH 212
Db 164 PGGPWCFYTTDPAVRFQSCGCIKSCREACVWCNGEYRGAVDRTESGRCQRDWLQHPHQH 223
QY 213 KFLPERYPDKGFDNRYCRNPDGQPRPMCYTLDPHTRWEYCAIKTCADNTMDTDVPLETT 272
Db 224 PFEPGKFLDQGLDDNYCRNPDGSESPWCFCTTDPOIEREFCDLPCGSEAOQRQA--TTV 281
QY 273 ECIQOGEYRGVTNTIWNIPCORWDSOYPEHDMTPTENFKDLRENYCRNPDGSESP 332
Db 282 SCFRGKGEYRGVTANTTTAGVPCQWRDQAIHQHRTPEKYACKDLRENFRCNPDGSEAP 341
QY 333 WCFTDPNIRVGYCSQIPNC--DMSHGQDCYRGNGKYNWGNLSQTRSGLTCSMDKNNMEDL 391
Db 342 WCFTLRPGCMRAAFVQIRRCTDDVRRPQDCYHAGEYRGVTSKTRKGVQCQWSAETPHK 401
QY 392 HRHIFWEPDASKLRENYCRNPDGDDAHGPMWCTGNPLIPMDYCPISRCGDDTTPTTV 447
Db 402 PQFTFTSPHAQLEENFCRNPDGSDSHGPMWCTMDPRTPTFDYCALRRCADDDPPSIL 457

RESULT 7
JC5061
macrophage-stimulating protein 1 precursor - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 31-Jan-1997 #sequence_revision 31-Jan-1997 #text_change 16-Jun-2000
C:Accession: JC5061
R:Ohshiro, K.; Iwama, A.; Matsuno, K.; Ezaki, T.; Sakamoto, O.; Hamaguchi, I.; Takasu
Biochem. Biophys. Res. Commun. 227, 273-280, 1996
A:Title: Molecular cloning of rat macrophage-stimulating protein and its involvement
A:Reference number: JC5061; MUID:97011126; PMID:8858136
A:Accession: JC5061
A:Molecule type: mRNA
A:Residues: 1-716 <OHS>
A:Cross-references: EMBL:X95096; NID:g1669718; PIDN:CAA64473.1; PID:g1669719
C:Complex: disulfide-bonded heterodimer of chains derived from the same precursor
C:Superfamily: hepatocyte growth factor; kringle homology; trypsin homology
C:Keywords: duplication; glycoprotein; growth factor; kringle
F:1-31/Domain: signal sequence #status predicted <SIG>
F:32-488,489-716/Product: macrophage-stimulating protein 1 #status predicted <MAT>
F:32-488/Domain: macrophage-stimulating protein 1 alpha chain #status predicted <ACH>
F:110-186/Domain: kringle homology <KR11>
F:191-268/Domain: kringle homology <KR12>
F:292-370/Domain: kringle homology <KR13>
F:379-457/Domain: kringle homology <KR14>
F:489-716/Domain: macrophage-stimulating protein 1 beta chain #status predicted <BCH>
F:489-709/Domain: trypsin homology <TRY>
F:72,305,620/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 43.98; Score 1144; DB 1; Length 716;
Best Local Similarity 46.18; Pred. No. 8.3e-74;
Matches 196; Conservative 70; Mismatches 141; Indels 18; Gaps 6;

```

```

QY 33 VNTADQCANRCTRNRKGLPFTCKAFVDFKARKQCLWFFPNSMSSGVYKKEFGHEFDLYENKD 92
Db 50 VADAEACARRC---GPLLDCRAFHYNSSHGCOLLPWTQHSPLRAQLHSSSLCDLFQKKD 105
QY 93 YIRNCIICKGRSYKGTIVTSIKSGIKCQPWSSMIPEHSFLPSSYRGKDLQENYCRNPRGE 152
Db 106 YVRTCIMDNGASYRGTVARTADGLPCQAWSRRFPNDHKYTTPTKNG--LEENFCRNPDGD 163
QY 153 EGGPWCFSTNPEVRYEVCIDIPOCSEVCMTCNGESYRGKGLMDHTESGKICQRDWDHQTPIRH 212
Db 164 PRGPWCYITNRSVRFQSCGCIKSCREAVCWNGEDYRGVDVTTESGRCQRDWLQHPHSH 223

```


Best Local Similarity 44.68; Pred. No. 3.2e-68; Matches 195; Conservative 65; Mismatches 137; Indels 40; Gaps 13;

Qy 25 ALKIKTKVNTADOCANCTRNKGLPFTCKAFVDFKARKQCLWPFNSMSGVKKKEFGHE 84
Db 16 SILSRQVAARSVEECAACEAETN--FCIRAFQYHSKQOQCVVMAENSKTSPARM--RD 71

Qy 85 FDLYENKDYIRNCIITGKRSYKGTVITKSGIKQCPWSSMIPHEHSFLPSSYRGKDLQEN 144
Db 72 VVLFKRIYLSCKTGTGNGKNGYRTSKTKSGVICQKWSVSPHIPKYSPEKFPPLAGLEEN 131

Qy 145 YCRNPRGEGGPWCFTSNPEVRYEVCIDIPOCSEVECMTCNGESYRGLMDHTESGKICQW 204
Db 132 YCRNPDNDEKGPWCYTTDPTETRFYCDIPEC--EDECMHCSGEHYEGKISKTMGSLGECQSW 190

Qy 205 DHQTPHRHKLPERYPDKGFDNYCRNPDGQPRPWCYITLDPHTRWECYCAIKTCADNTWND 264
Db 191 GSOSPHAGLYLSPKPNKLNMYCRNPDGEPWPWCFTTDPNKRWFCDIPRC--TTPPP 248

Qy 265 TDVPLETTECQOGGEGYRGVTWNTWINGIPQORWDSQYPHEHDMTPENFKCKDLRENYCR 324
Db 249 TSGP--TYQCLKGRGENYRGVTWNTWINGIPQORWDSQYPHEHDMTPENFKCKDLRENYCR 306

Qy 325 NPDGSESPWCFTTDPNIRVGVCSQIPNCDMS-----H-----GQDCYRGNGKN 367
Db 307 NPDGSESPWCFTTDPNIRVGVCSQIPNCDMS-----H-----GQDCYRGNGKN 367

Qy 368 YGNSLQTSRSLGTCMMDKNNEDLHRHFWB-----PDASKLNENYCRNPDGQPRPWCYITLDPHTRWECYCAIKTCADNTWND 421
Db 366 YRGTSSTITGRKQSWVSWTP--HRH---EKTQCNFFNAG--LTWNYCRNPDG--KSPWC 418

Qy 422 YTGPNLPDWCPISRC 438
Db 419 YTTDPRVRWEYCNLKKC 435

RESULT 10
PLMS
A:Title: Characterization of the cDNA coding for mouse plasminogen and localization of the plasminogen gene.
A:Accession: A38514
A:Reference number: A38514; MUID:91184812; PMID:2081600
A:Molecule type: mRNA
A:Residues: 1-812 <DE>
A:Cross-references: GB:J04766; NID:g200402; PIDN:AAA50168.1; PID:g200403
R:Lijnen, H.R.; van Hoef, B.; Beelen, V.; Collen, D.
Eur. J. Biochem. 224, 863-871, 1994
A:Title: Characterization of the murine plasma fibrinolytic system.
A:Reference number: S48202; MUID:95010076; PMID:7523120
A:Accession: S48202
A:Molecule type: protein
A:Residues: 20-25
A:Accession: S48203
A:Molecule type: protein
A:Residues: 22-27
C:Comment: Plasminogen is synthesized by the kidney and is present in plasma and many other tissues. Plasminogen is converted into plasmin by plasminogen activators, both plasminogen activators and plasminogen activators, both plasminogen activators and plasminogen activators.
C:Comment: Plasminogen is synthesized by the kidney and is present in plasma and many other tissues. Plasminogen is converted into plasmin by plasminogen activators, both plasminogen activators and plasminogen activators, both plasminogen activators and plasminogen activators.
C:Comment: Plasminogen is synthesized by the kidney and is present in plasma and many other tissues. Plasminogen is converted into plasmin by plasminogen activators, both plasminogen activators and plasminogen activators, both plasminogen activators and plasminogen activators.

F:1-96/Domain: plasminogen-related protein precursor homology <PLPH>
F:1-19/Domain: signal sequence #status predicted <SIG>
F:1-812/Product: plasminogen #status predicted <PRO>
F:20-96/Domain: activation peptide #status predicted <APT>
F:79-466/Product: angiotensin #status predicted <AST>
F:97-581,582-812/Product: plasmin #status predicted <MAT>
F:97-581/Domain: chain A #status predicted <ACH>
F:103-181/Domain: kringle homology <KR1>
F:185-262/Domain: kringle homology <KR2>
F:275-352/Domain: kringle homology <KR3>
F:377-454/Domain: kringle homology <KR4>
F:481-560/Domain: kringle homology <KR5>
F:582-812/Domain: chain B #status predicted <BCH>
F:582-805/Domain: trypsin homology <TRY>
F:49-73,53-61,103-181,124-164,152-176,185-262,188-316,206-245,234-257,275-352,296-335
bonds: #status predicted
F:78-79/Cleavage site: Glu-Asn (stromelysin 1) #status predicted
F:136,308/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:466-467/Cleavage site: Thr-Val (stromelysin 1) #status predicted
F:581-582/Cleavage site: Arg-Val (plasminogen activator) #status experimental
F:624,667,762/Active site: His, Asp, Ser #status predicted

Query Match 40.7%; Score 1059; DB 1; Length 812;
Best Local Similarity 43.3%; Pred. No. 1e-67; Mismatches 145; Indels 38; Gaps 12;
Matches 191; Conservative 67

Qy 30 TKKVNTA---DQCANRCTRNKGLPFTCKAFVDFKARKQCLWPFNSMSGVKKKEFGHEFD 86
Db 37 TKQLAAGVSDCLAKCEGE--TDFVCRSFQYHSKEQCVMAENSKTSSIIRM--RDVI 92

Qy 87 LYENKDYIRNCIITGKRSYKGTVITKSGIKQCPWSSMIPHEHSFLPSSYRGKDLQENYC 146
Db 93 LFEKRVYLSCKTGTGNGYRGVTWNTWINGIPQORWDSQYPHEHDMTPENFKCKDLRENYC 152

Qy 147 RNPRGEGGPWCFTSNPEVRYEVCIDIPOCSEVECMTCNGESYRGLMDHTESGKICQW 206
Db 153 RNPDNDEKGPWCYTTDPTETRFYCDIPEC--EDECMHCSGEHYEGKISKTMGSLGECQSW 211

Qy 207 QTPHRHKLPERYPDKGFDNYCRNPDGQPRPWCYITLDPHTRWECYCAIKTCADNTWNDTD 266
Db 212 QSPHAGYIPAKFPKPNKLNMYCRNPDGEPWPWCFTTDPNKRWFCDIPRC--TTPPP 267

Qy 267 VPLETTECQOGGEGYRGVTWNTWINGIPQORWDSQYPHEHDMTPENFKCKDLRENYCRNP 326
Db 268 PPSPTYQCLKGRGENYRGVTWNTWINGIPQORWDSQYPHEHDMTPENFKCKDLRENYC 327

Qy 327 DGSSESPWCFTTDPNIRVGVCSQIPNCDMSHG-----ODCYRGNGKNYM 369
Db 328 DGETAPWCYTTDPTETRFYCDIPEC--EDECMHCSGEHYEGKISKTMGSLGECQSW 386

Qy 370 GNLQTSRSLGTCMMDKNNEDLHRHFWB---PDASKLNENYCRNPDGQPRPWCYITGNP 426
Db 387 GTSSTITGRKQSWVSWTP--HRH---EKTQCNFFNAG--LTWNYCRNPDG--KSPWC 442

Qy 427 LIPDWCPISRCGEGTTP 447
Db 443 SVRWEYCNLKR--SETGGSVV 462

RESULT 11
I46260
A:Title: The recurring evolution of Lp(a): Insights from cloning of hedgehog apolipoprotein C3.
A:Reference number: I46260; MUID:96025778; PMID:7592597
A:Accession: I46260
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-810 <LAW>

A:Cross-references: EMBL:U33171; NID:g1046360; PID:g1046361
C:Superfamily: plasmin; kringle homology; plasminogen-related protein precursor homology
C:Keywords: hydrolase; serine proteinase
F:1-96/Domain: plasminogen-related protein precursor homology <PLPH>
F:103-181/Domain: kringle homology <KR1>
F:185-262/Domain: kringle homology <KR2>
F:275-352/Domain: kringle homology <KR3>
F:379-456/Domain: kringle homology <KR4>
F:482-561/Domain: kringle homology <KR5>
F:582-803/Domain: trypsin homology <TRY>

Query Match 40.5%; Score 1053.5; DB 2; Length 810;
Best Local Similarity 44.3%; Pred. No. 2.6e-67;
Matches 194; Conservative 63; Mismatches 138; Indels 43; Gaps 13;

QY 30 TKK--VNTADOCANRCTRNGKGLPTCTCAFAVDKARKQCLWFFPNSMSSGYKKEGHEFD 86
DB 37 TKQLSVGSTEECAVKC--EKETSFICRSFYHSKEQOCVMAENRKSIVLRM--RDVI 92

QY 87 LYENKDYIRNCIIGKRSYKGTIVTSITKSGIKQCPWSSMIPHEHSFLPSSYRGKDLQENYC 146
DB 93 LFEKKMYLSECKVNGKYYRGTVSKTGLTCQKWSAETPHKPRESPDENPSEGLDQNYC 152

QY 147 RNPREGGPGWCFSTNPEVRYEVCIDIPQCSVECMTCNGESYRGLMDHTESGKICQWDH 206
DB 153 RNPDPKPGKPCWCTMDPEVRYEVCETIQ--EDECHMCSGQNVGKISRTMSGLEQCPWDS 211

QY 207 QTPHRHKLFLPERYPDKGFDNDCRNPDGPRPCWCTLDPHTRWEYCAIKTCADNTMDTD 266
DB 212 QIPPHGFIPTSPKPNLKNKNCNPDGPRPCWCTMDRKNRWEYCDIPRCT-----TP 265

QY 267 VPLE--TTECIQOGGEGYRGTVNTIWNIGIPQWDSQYRPHEDMTPEKCKDLRENCR 324
DB 266 PPSGGTYQCLMNGEYHOGVAVTVSGLTQCRWGEQSPHRHRTPEYPCNLDENYCR 325

QY 325 NPDGSESPWCFSTDNIRVGYCSQIPNC-----DMSHG-----QDCYRGNG 365
DB 326 NPDGEPAPWCFSTNSVRWEFC-KIPDCVSSASSETSHPDAPVTPPEQPPVVOECYQNG 384

QY 366 KNYGNLSQTRSLGTCMSMDKNEHLHFIWE--PDASKLNENYCRNPDADDHGPWCY 422
DB 385 QTYRGTSSTITGKKOPWTSMRP--HRHSKTPENYPDAD-LTMNCRNPDGD-KGPWCY 440

QY 423 TGNPLIPWDYCPISRCGE 440
DB 441 TTDPSVRWEFCNLKCKSG 458

RESULT 12
B30848
plasmin (EC 3.4.21.7) precursor - rhesus macaque
C:Species: Macaca mulatta (rhesus macaque)
C:Date: 31-Mar-1989 #sequence_revision 31-Mar-1989 #text_change 22-Jun-1999
C:Accession: B32869; B30848
R:Tominson, J.E.; McLean, J.W.; Lawn, R.M.
J. Biol. Chem. 264, 5957-5965, 1989
A:Title: Rhesus monkey apolipoprotein(a). Sequence, evolution, and sites of synthesis.
A:Reference number: A32869; MUID:89174660; PMID:2925643
A:Accession: B32869
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-810 <TOM>
A:Cross-references: GB:J04697; NID:g342272; PIDN:AAA36901.1; PID:g342273
C:Superfamily: plasmin; kringle homology; plasminogen-related protein precursor homology
C:Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase
F:1-96/Domain: plasminogen-related protein precursor homology <PLPH>
F:103-181/Domain: signal sequence #status predicted <SIG>
F:185-262/Domain: kringle homology <KR1>
F:275-352/Domain: kringle homology <KR2>
F:379-456/Domain: kringle homology <KR3>
F:481-560/Domain: kringle homology <KR4>
F:581-803/Domain: trypsin homology <TRY>

F:49-73,53-61,103-181,124-164,152-176,185-262,188-316,206-245,234-257,275-352,296-335
bonds: #status predicted
F:622,665,760/Active site: His, Asp, Ser #status predicted

Query Match 40.1%; Score 1044.5; DB 2; Length 810;
Best Local Similarity 42.6%; Pred. No. 1.1e-66;
Matches 188; Conservative 74; Mismatches 128; Indels 51; Gaps 12;

QY 30 TKK--VNTADOCANRCTRNGKGLPTCTCAFAVDKARKQCLWFFPNSMSSGYKKEGHEFD 86
DB 37 TKQLGAGSIECAKCEEE--EFTCRSFQYHSKEQOCVMAENRKSIVLRM--RDVV 92

QY 87 LYENKDYIRNCIIGKRSYKGTIVTSITKSGIKQCPWSSMIPHEHSFLPSSYRGKDLQENYC 146
DB 93 LFEKKVYLSECKTGNKYNRGTSKTRTGITCQKWSSTSPHRPTSPATHPSEGLEENYC 152

QY 147 RNPREGGPGWCFSTNPEVRYEVCIDIPQCSVECMTCNGESYRGLMDHTESGKICQWDH 206
DB 153 RNPNDGGPGWCTYTDPEEREDYCDIPEC--EDECHMCSGENYDGKISKTMGLEQCAWDS 211

QY 207 QTPHRHKLFLPERYPDKGFDNDCRNPDGPRPCWCTLDPHTRWEYCAIKTCADNTMDTD 266
DB 212 QSPHAHGYIPSKFKNLKNKNCNPDGPRPCWCTTDPNKRKWLCDIPRCT-----TP 265

QY 267 VPLE--TTECIQOGGEGYRGTVNTIWNIGIPQWDSQYRPHEDMTPEKCKDLRENCR 324
DB 266 PPSGGTYQCLKGTGENYRGDVAVTVSGTSGHSAQTPHTHNRTPENFPCNLDENYCR 325

QY 325 NPDGSESPWCFSTDNIRVGYCSQIPNCDSH-----GDYCRNGKN 367
DB 326 NPDGKAPWCTYTSNQVRWEYC-KIPSCSSPVSTEPDPTAPPELTPVVOECYHGDQS 384

QY 368 YMGNLISQTRSLGTCMSW-----DKNMDLHRHFIWEPPDASKLNENYCRNPDADDHGP 419
DB 365 YRGTSSTITGKKQCSWSSMTPHWHEKTPENF-----PNAG-LTMNCRNPDAD-KGP 435

QY 420 WCYTNGLIPWDYCPISRCGE 440
DB 436 WCFTTDPSPVRWEYCNLKKCSG 456

RESULT 13

PLHU

plasmin (EC 3.4.21.7) precursor [validated] - human
N:Alternate names: plasminogen precursor [misnomer]
N:Contains: angiotatin; microplasmin; plasminogen
C:Species: Homo sapiens (man)

C:Date: 24-Apr-1984 #sequence_revision 02-Dec-1994 #text_change 15-Sep-2000
C:Accession: A35229; I52242; A26646; I62738; I84609; S03735; A00929; A04627; A04625;
J.Petersen, T.E.; Martzen, M.R.; Ichinose, A.; Davie, E.W.
J. Biol. Chem. 265, 6104-6111, 1990
A:Title: Characterization of the gene for human plasminogen, a key proenzyme in the f
A:Reference number: A35229; MUID:90202879; PMID:2318848

A:Accession: A35229

A:Molecule type: DNA

A:Residues: 1-810 <PE>

A:Cross-references: GB:J05286; GB:M34276; NID:g190064; PIDN:AAA60113.1; PID:g387026
A:Experimental source: leukocyte; lung fibroblast
R:Maigaretta, N.; Bruno, L.; Pontoglio, M.; Candiani, G.; Meroni, G.; Ottolenghi, S.;
Biochem. Biophys. Res. Commun. 173, 1013-1018, 1990
A:Title: Definition of the transcription initiation site of human plasminogen gene in
A:Reference number: I52242; MUID:91097523; PMID:2268308

A:Accession: I52242

A:Status: translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-16 <MALI>

A:Cross-references: GB:M62890; NID:g190092; PIDN:AAA36454.1; PID:g553613

R:Forssgren, M.; Raden, B.; Israelsson, M.; Larsson, K.; Heden, L.O.

FEBS Lett. 213, 254-260, 1987

A:Title: Molecular cloning and characterization of a full-length cDNA clone for human
A:Reference number: A26646; MUID:87162490; PMID:3030813

A:Accession: A26646

A:Molecule type: mRNA

- A:Residues: 1-471,'D',473-810 <FOR>
A:Cross-references: GB:X05199; NID:g35530; PIDN:CAA28831.1; PID:g35531
A:Experimental source: liver
R:Malinowski, D.P.; Sadler, J.E.; Davie, E.W.
Biochemistry 23, 4243-4250, 1984
A:Title: Characterization of a complementary deoxyribonucleic acid coding for human and
A:Reference number: I45961; MUID:85023311; PMID:6148961
A:Accession: I62738
A>Status: translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 292-471,'D',473-810 <MAL2>
A:Cross-references: GB:K02922; NID:g190112; PIDN:AAA60124.1; PID:g387031
A:Accession: I84609
A>Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 367-419 <MAL3>
A:Cross-references: GB:K02921; NID:g190110; PIDN:AAA60123.1; PID:g190111
R:Brünsholz, R.A.; Lerch, P.G.; Schaller, J.; Rickli, E.E.; Lergier, W.; Manneberg, M.;
Eur. J. Biochem. 114, 463-470, 1981
A:Title: Comparison of the primary structure of the N-terminal CNBR fragments of human,
A:Reference number: S03735; MUID:81212097; PMID:7238497
A:Accession: S03735
A:Molecule type: protein
A:Residues: 20-71,'E',73-85,87-106,'D',108-360,'E',362-810 <SOT>
R:Wlman, B.
A:Residues: 20-71,'E',73-85,87-106,'D',108-360,'E',362-810 <SOT>
R:Wlman, B.
A:Title: Primary structure of the B-chain of human plasmin.
A:Reference number: A04627; MUID:7725245; PMID:142009
A:Accession: A04627
A:Molecule type: protein
A:Residues: 581-810 <WIL>
R:Wlman, B.; Wallen, P.
Eur. J. Biochem. 50, 489-494, 1975
A:Title: Structural relationship between "glutamic acid" and "lysine" forms of human plasminogen
A:Reference number: A04625; MUID:75093329; PMID:122932
A:Accession: A04625
A:Molecule type: protein
A:Residues: 20-50,'Q',51-71,'E',73-85,87-100 <W12>
R:Wlman, B.; Wallen, P.
Eur. J. Biochem. 58, 539-547, 1975
A:Title: Amino-acid sequence of the cyanogen-bromide fragment from human plasminogen the
A:Reference number: A04626; MUID:76043692; PMID:126863
A:Accession: A04626
A:Molecule type: protein
A:Residues: 483-507,'E',509-604 <W13>
R:Robbins, K.C.; Bernabe, P.; Arzadon, L.; Summaria, L.
J. Biol. Chem. 248, 1631-1633, 1973
A:Title: The primary structure of human plasminogen. II. The histidine loop of human plasminogen
A:Reference number: A92125; MUID:73149248; PMID:4694729
A:Contents: annotation: active site
R:Groskopf, W.R.; Summaria, L.; Robbins, K.C.
J. Biol. Chem. 244, 3590-3597, 1969
A:Title: Studies on the active center of human plasmin. Partial amino acid sequence of a
A:Reference number: A92048; MUID:69234739; PMID:4240117
A:Contents: annotation: active site
R:Trexler, M.; Vali, Z.; Pathy, L.
J. Biol. Chem. 257, 7401-7406, 1982
A:Title: Structure of the omega-aminocarboxylic acid-binding sites of human plasminogen.
A:Reference number: A92382; MUID:82213905; PMID:6919539
A:Contents: annotation: omega-aminocarboxylic acid binding sites
R:Vali, Z.; Pathy, L.
J. Biol. Chem. 259, 13690-13694, 1984
A:Title: The fibrin-binding site of human plasminogen. Arginines 32 and 34 are essential
A:Reference number: A92458; MUID:85054794; PMID:6094526
A:Contents: annotation: fibrin binding site; omega-aminocarboxylic acid binding site
R:Cao, Y.; Ji, R.W.; Davidson, D.; Schaller, J.; Marti, D.; Soehnle, S.; McCance, S.G.;
J. Biol. Chem. 271, 29461-29467, 1996
A:Title: Kringle domains of human angiotensin. Characterization of the anti-proliferative
A:Reference number: A58811; MUID:97067211; PMID:8910613
A:Contents: annotation
R:Lijnen, H.R.; Ugwu, F.; Bini, A.; Collen, D.
Biochemistry 37, 4699-4702, 1998
A:Title: Generation of an angiotensin-like fragment from plasminogen by stromelysin-1
A:Reference number: A58812; MUID:9548733; PMID:9548733
A:Contents: annotation
R:Tulinsky, A.; Mulichak, A.M.
submitted to the Brookhaven Protein Data Bank, July 1991
A:Reference number: A51341; PDB:1PK4
A:Contents: annotation; X-ray crystallography, 1.9 angstroms, residues 376-454
R:Tulinsky, A.; Wu, T.P.
submitted to the Brookhaven Protein Data Bank, July 1991
A:Reference number: A51488; PDB:2PK4
A:Contents: annotation; X-ray crystallography, 2.25 angstroms, residues 375-454
R:Wu, T.P.; Tulinsky, A.
submitted to the Brookhaven Protein Data Bank, August 1993
A:Reference number: A51911; PDB:1PKR
A:Contents: annotation; X-ray crystallography, 2.48 angstroms, residues 102-181
R:Padmanabhan, K.; Tulinsky, A.
submitted to the Brookhaven Protein Data Bank, April 1994
A:Reference number: A52408; PDB:1PMK
A:Contents: annotation; X-ray crystallography, 2.25 angstroms, residues 377-454
R:Tulinsky, A.; Mathews, I.I.
submitted to the Brookhaven Protein Data Bank, December 1995
A:Reference number: A65244; PDB:1CEA
A:Contents: annotation; X-ray crystallography, 2.1 angstroms, residues 102-181
R:Tulinsky, A.; Mathews, I.I.
submitted to the Brookhaven Protein Data Bank, December 1995
A:Reference number: A65245; PDB:1CEB
A:Contents: annotation; X-ray crystallography, 2.1 angstroms, residues 102-181
R:Mulichak, A.M.; Tulinsky, A.; Ravichandran, K.G.
Biochemistry 30, 10576-10588, 1991
A:Title: Crystal and molecular structure of human plasminogen kringle 4 refined at 1.
A:Reference number: A58819; MUID:92031502; PMID:1657148
A:Contents: annotation
R:Wu, T.P.; Padmanabhan, K.; Tulinsky, A.; Mulichak, A.M.
Biochemistry 30, 10589-10594, 1991
A:Title: The refined structure of the epsilon-aminocaproic acid complex of human plasminogen
A:Reference number: A58818; MUID:92031503; PMID:1657149
A:Contents: annotation
R:de Vos, A.M.; Uitsch, M.H.; Kelley, R.F.; Padmanabhan, K.; Tulinsky, A.; Westbrook, J.
Biochemistry 31, 270-279, 1992
A:Title: Crystal structure of the kringle 2 domain of tissue plasminogen activator at
A:Reference number: A39483; MUID:92118803; PMID:1310033
A:Contents: annotation; X-ray crystallography, 2.4 angstroms
R:Stec, B.; Teeter, M.M.; Whitlow, M.; Yamano, A.
submitted to the Brookhaven Protein Data Bank, June 1995
A:Reference number: A65980; PDB:1KRN
A:Contents: annotation; X-ray crystallography, 1.67 angstroms, residues 376-454
R:Rejante, M.; Llinas, M.
submitted to the Brookhaven Protein Data Bank, August 1996
A:Reference number: A65803; PDB:1HPJ
A:Contents: annotation; conformation by (1)H-NMR, residues 103-181
R:Rejante, M.; Llinas, M.
submitted to the Brookhaven Protein Data Bank, August 1996
A:Reference number: A65804; PDB:1HPK
A:Contents: annotation; conformation by (1)H-NMR, residues 103-181
R:Rejante, M.R.; Llinas, M.
Eur. J. Biochem. 221, 927-937, 1994
A:Title: (1)H-NMR assignments and secondary structure of human plasminogen kringle 1.
A:Reference number: S43645; MUID:94237157; PMID:8181475
A:Contents: annotation; conformation by (1)H-NMR, residues 96-184
R:Rejante, M.R.; Llinas, M.
Eur. J. Biochem. 221, 939-949, 1994
A:Title: Solution structure of the epsilon-aminohexanoic acid complex of human plasminogen
A:Reference number: A58817; MUID:94237158; PMID:8181476
A:Contents: annotation; conformation by (1)H-NMR
C:Comment: plasminogen is synthesized by the kidney and is present in plasma and many
C:Comment: plasminogen is converted to plasmin by plasminogen activators (see PIR:UKH
d PIR:FGHUGB).

Search completed: April 25, 2003, 15:52:36
Job time : 51 secs

```
Db 158 YCRNPDNDENGFWCYTTDPDKRYDYCDIPEC-EDKCMHCSGENYEGKIATKTMGRDCQAW 216
Qy 205 DHQTPHRHKFLPERYPDGFDDNYCRNPDGQPRWCYTLDPHTRWEYCAIKTCADNTWMD 264
Db 217 DSQSPHAGYIPSPFNKLNLMXNYCRNPDGEPWCFTTDPQKRWEFCDIPTCT----- 270
Qy 265 TDVPLE--TTTCIOGQGGYRGTVNTIWNIGPCQWDSQYPHEHDMTPENPKCKDLRENY 322
Db 271 TPPSSGPKYQCLGTYGNYGCTVAVTESGHTCQWSEQTPHKHNRTPENFPCKNLENY 330
Qy 323 CRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSH-----GQDCYRNG 365
Db 331 CRNPNGERAPWCYTINSRWEYCT-IPSCSSPLSTERMDVPVPEQTPVQDCYHNG 389
Qy 366 KNYGNLSQTRSGLTCSMDKNMEDLHRHFEW----PDASKLNENYCRNPDDDAHGFWCY 422
Db 390 QSYRCTSTTITGRKQCSWSSWTP--HRHLKTPENYPNAG-LTMNYCRNPDAK-KSPWCY 445
Qy 423 TGNPLIPWDYCPISRC 438
Db 446 TTDPRVRWEFCNLKKC 461

RESULT 15
Tl8518
apolipoprotein(a) - western European hedgehog (fragment)
C:Species: Erinaceus europaeus (western European hedgehog)
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 07-Dec-1999
C:Accession: Tl8518
R:Lawn, R.M.; Boonmark, N.W.; Schwartz, K.; Lindahl, G.E.; Wade, D.P.; Byrne, C.D.; Fong
J. Biol. Chem. 270, 24004-24009, 1995
A:Title: The recurring evolution of Lp(a): Insights from cloning of hedgehog apolipoprotein
A:Reference number: 146259; MUID:96025778; PMID:7592597
A:Accession: Tl8518
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-2869 <LAW>
A:Cross-references: EMBL:U33170; NID:gl046358; PID:gl046359; PIDN:AAC48522.1
A:Experimental source: liver
C:Comment: The lipoprotein Lp(a), a major inherited risk factor for atherosclerosis, contains
ent apolipoprotein(a).

Query Match 36.2%; Score 943.5; DB 2; Length 2869;
*Best Local Similarity 45.6%; Pred. No. 6.4e-59;
Matches 173; Conservative 51; Mismatches 126; Indels 29; Gaps 10;

Qy 89 ENKDYIRNCIIGKRSYKGTVSITKSGIKCQWPSSMIPHSFPLSPSYRGKDLQENYCRN 148
Db 2490 EPLDLVDQCLEGTGENYGNNAVTAAGHTCQWRQSPHSRTPENYPTKNLVGNYCRN 2549
Qy 149 PRGEEGPPWCFTSNPEVRYEVCIDIQCSE-----VECMTCNGESYRGLMDHTES 197
Db 2550 PDGEV-APWCYTNSAMRWEYCSIPACESPTPTTEHLVPEQCLEGNGENYQGNNAITVS 2608
Qy 198 GKICORWDHQTPIRHKFLPERYPDKGFDNVCNRPDQGPWCYTLDPHTRWEYCAIKTC 257
Db 2609 GQPCQGWKQKTPHREYTPENYPSNLFNGYCRNPDGAIAPWCYTNTSAVRWEYCSIPTC 2668
Qy 258 ADNTMNDTDVPLETTECIQGGEGYRGTVNTIWNIGPCQWDSQYPHEHDMTPENPKCKD 317
Db 2669 -ESSSPTEPMIIPQCLEGTGENYRGNAVTVSGHTCQWRQSPHSRTPENYPTKN 2727
Qy 318 LRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSH-----GQD-CYRGNKNYM 369
Db 2728 LFGNYCRNPDGEVAPWCYTNSAVRWEYC-KIPSCNSSSPTEPSTSDQCLKNGESYQ 2786
Qy 370 GNLSQTRSGLTCSMDKNMEDLHRHFEWPD---ASKLNENYCRNPDDDAHGFWCYTGNP 426
Db 2787 GNISVTSGYTCQWRRE--QTPHRAA-RTPDNYCKNLVGNYCRNPDGEV-APWCYTNS 2842
Qy 427 LIPWDYCPISRCGDTTPT 445
Db 2843 AVRWEYCSIPTCESPTPT 2861
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 25, 2003, 15:45:29 ; Search time 25 Seconds
(without alignments)
741.597 Million cell updates/sec

Title: US-09-674-377b-1

Perfect score: 2604

Sequence: 1 ERRRRNTIHEFKSAKTTLI.....IPWDYCPISRCGDTTPTIV 447

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match %	Score	Length	DB ID	Description
1	2601	99.9	728	1	HGF_HUMAN
2	2408	92.5	728	1	HGF_MOUSE
3	2401	92.2	728	1	HGF_RAT
4	1180.5	45.3	711	1	HGFL_HUMAN
5	1125	43.2	716	1	HGFL_MOUSE
6	1066	40.9	790	1	PLMN_PIG
7	1059	40.7	812	1	PLMN_MOUSE
8	1053.5	40.5	810	1	PLMN_ERIEU
9	1044.5	40.1	810	1	PLMN_MACMU
10	1042	40.0	810	1	PLMN_HUMAN
11	1038	39.9	812	1	PLMN_BOVIN
12	879	33.8	4548	1	APOA_HUMAN
13	843.5	32.4	1420	1	APOA_MACMU
14	408.5	15.7	325	1	PLMN_PETMA
15	357.5	13.7	625	1	THRB_BOVIN
16	353	13.6	169	1	PLMN_RAT
17	352.5	13.5	618	1	THRB_MOUSE
18	350	13.4	622	1	THRB_HUMAN
19	344	13.2	617	1	THRB_RAT
20	307.5	11.8	559	1	TPA_RAT
21	303	11.6	562	1	TPA_HUMAN
22	303	11.6	566	1	TPA_BOVIN
23	297.5	11.4	559	1	TPA_MOUSE
24	253	9.7	338	1	PLMN_HORSE
25	241	9.3	343	1	PLMN_SHEEP
26	237	9.1	333	1	PLMN_CANFA
27	230	8.8	943	1	ROR2_HUMAN
28	228.5	8.8	603	1	FA12_CAVPO
29	223	8.6	944	1	ROR2_MOUSE
30	220	8.4	937	1	ROR1_MOUSE
31	217	8.3	937	1	ROR1_HUMAN
32	204	7.8	655	1	HGFA_HUMAN
33	198	7.6	653	1	HGFA_MOUSE

34 195 7.5 615 1 FA12_HUMAN p00748 homo sapien
35 189 7.3 593 1 FA12_BOVIN p98140 bos taurus
36 187 7.2 473 1 KREM_MOUSE Q95n43 mus musculus
37 187 7.2 473 1 KREM_RAT Q92484 rattus norv
38 180.5 6.9 475 1 KREM_HUMAN Q96nu8 homo sapien
39 158.5 6.1 433 1 UROK_BOVIN Q05589 bos taurus
40 158 6.1 442 1 UROK_PIG P04185 sus scrofa
41 156 6.0 477 1 URT1_DESRO P98119 desmodus ro
42 152.5 5.9 875 1 NETR_HUMAN P56730 homo sapien
43 151.5 5.8 431 1 UROK_HUMAN P00749 homo sapien
44 151 5.8 432 1 UROK_RAT P29598 rattus norv
45 151 5.8 477 1 URT2_DESRO P15638 desmodus ro

ALIGNMENTS

RESULT 1
HGF_HUMAN
ID HGF_HUMAN STANDARD; PRT; 728 AA:
AC P14210; Q9UDU6; Q9BYL9;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hepatocyte growth factor precursor (Scatter factor) (SF)
DE (Hepatopoietin A).
GN HGF OR HPTA.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91340155; PubMed=1831432;
RA Seki T., Hagiya M., Shimonishi M., Nakamura T., Shimizu S.;
RT "Organization of the human hepatocyte growth factor-encoding gene.";
RL Gene 102:213-219(1991).
RN [2]
RP SEQUENCE FROM N.A.
RX TISSUE=Placenta;
RA MEDLINE=89392017; PubMed=2528952;
RA Miyazawa K., Tsubouchi H., Naka D., Takahashi K., Okigaki M.,
RA Arakaki N., Nakayama H., Hirono S., Sakiyama O., Takahashi K.,
RA Gohda E., Daikuhara Y., Kitamura N.;
RT "Molecular cloning and sequence analysis of cDNA for human hepatocyte
growth factor.";
RL Biochem. Biophys. Res. Commun. 163:967-973(1989).
RN [3]
RP SEQUENCE FROM N.A.
RX TISSUE=Leukocyte;
RA MEDLINE=91025062; PubMed=2145836;
RA Seki T., Ihara I., Sugimura A., Shimonishi M., Nishizawa T.,
RA Asami O., Hagiya M., Nakamura T., Shimizu S.;
RT "Isolation and expression of cDNA for different forms of hepatocyte
growth factor from human leukocyte.";
RL Biochem. Biophys. Res. Commun. 172:321-327(1990).
RN [4]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 55-73 AND 495-520.
RX TISSUE=Liver;
RA MEDLINE=90066676; PubMed=2531289;
RA Nakamura T., Nishizawa T., Hagiya M., Seki T., Shimonishi M.,
RA Sugimura A., Tashiro K., Shimizu S.;
RT "Molecular cloning and expression of human hepatocyte growth factor.";
RL Nature 342:440-443(1989).
RN [5]
RP SEQUENCE FROM N.A.
RX TISSUE=Embryonic fibroblast;
RA MEDLINE=91334393; PubMed=1831266;
RA Weidner K.M., Arakaki N., Hartmann G., Vandekerckhove J., Weingart S.,
RA Rieder H., Fonatsch C., Tsubouchi H., Hishida T., Daikuhara Y.,
RA Birchnieker W.;
RT "Evidence for the identity of human scatter factor and human
hepatocyte growth factor.";

Proc. Natl. Acad. Sci. U.S.A. 88:7001-7005(1991).

[6] SEQUENCE FROM N.A.

RA Courtney L., Elliot G., Angell S.;

RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.

[7]

RN SEQUENCE OF 249-695 FROM N.A.

RX MEDLINE=91369928; PubMed=1832556;

RA Miyazawa K., Kitamura A., Kitamura N.;

RL "Structural organization and the transcription initiation site of the

RT human hepatocyte growth factor gene.";

RL Biochemistry 30:9170-9176(1991).

[8]

RN SIGNAL SEQUENCE CLEAVAGE SITE.

RP MEDLINE=91207365; PubMed=1826837;

RA Yoshizawa Y., Arakaki N., Naka D., Takahashi K., Hirose S., Kondo J.,

RA Nakayama H., Gohda E., Kitamura N., Tsubouchi H., Ishii T.,

RA Hishida T., Daikuhara Y.;

RT "Identification of the N-terminal residue of the heavy chain of both

RT native and recombinant human hepatocyte growth factor.";

RL Biochem. Biophys. Res. Commun. 175:660-667(1991).

[9]

RN CARBOHYDRATE-LINKAGE SITE THR-476.

RP MEDLINE=93129192; PubMed=1482348;

RA Shimizu N., Hara H., Sogabe T., Sakai H., Ihara I., Inoue H.,

RA Nakamura T., Shimizu S.;

RT "Hepatocyte growth factor is linked by O-glycosylated oligosaccharide

RT on the alpha chain.";

RL Biochem. Biophys. Res. Commun. 189:1329-1335(1992).

[10]

RN MUTAGENESIS.

RP MEDLINE=92331602; PubMed=1321034;

RA Lokker N.A., Mark M.R., Luis E.A., Bennett G.L., Robbins K.A.,

RA Baker J.B., Godowski P.J.;

RT "Structure-function analysis of hepatocyte growth factor:

RT Identification of variants that lack mitogenic activity yet retain

RT high affinity receptor binding.";

RL EMBO J. 11:2503-2510(1992).

[11]

RN STRUCTURE BY NMR OF 31-127.

RP MEDLINE=98154323; PubMed=9493272;

RA Zhou H., Mazzulla M.J., Kaufman J.D., Stahl S.J., Wingfield P.T.,

RA Rubin J.S., Bottaro D.P., Byrd R.A.;

RT "The solution structure of the N-terminal domain of hepatocyte growth

RT factor reveals a potential heparin-binding site.";

RL Structure 6:109-116(1998).

[12]

RN X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 35-210.

RP MEDLINE=99036858; PubMed=9817840;

RA Uitsch M., Lokker N.A., Godowski P.J., de Vos A.M.;

RT "Crystal structure of the NK1 fragment of human hepatocyte growth

RT factor at 2.0-A resolution.";

RL Structure 6:1383-1393(1998).

CC -1- FUNCTION: HGF IS A POTENT MITOGEN FOR MATURE PARENCHYMAL

CC HEPATOCYTE CELLS, SEEMS TO BE AN HEPATOTROPIC FACTOR, AND ACTS

CC AS GROWTH FACTOR FOR A BROAD SPECTRUM OF TISSUES AND CELL TYPES.

CC IT HAS NO DETECTABLE PROTEASE ACTIVITY.

CC -1- SUBUNIT: DIMER OF AN ALPHA CHAIN AND A BETA CHAIN LINKED BY A

CC DISULFIDE BOND.

CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. PLASMINOGEN SUBFAMILY.

CC -1- SIMILARITY: CONTAINS 4 KRINGLE DOMAINS.

CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC

DR EMBL; D90320; BAA14348.1; JOINED.

DR EMBL; D90322; BAA14348.1; JOINED.

DR EMBL; D90323; BAA14348.1; JOINED.

DR EMBL; D90324; BAA14348.1; JOINED.

DR EMBL; D90325; BAA14348.1; JOINED.

DR EMBL; D90326; BAA14348.1; JOINED.

DR EMBL; D90327; BAA14348.1; JOINED.

DR EMBL; D90328; BAA14348.1; JOINED.

DR EMBL; D90329; BAA14348.1; JOINED.

DR EMBL; D90330; BAA14348.1; JOINED.

DR EMBL; D90331; BAA14348.1; JOINED.

DR EMBL; D90332; BAA14348.1; JOINED.

DR EMBL; D90333; BAA14348.1; JOINED.

DR EMBL; M29145; AAA52650.1; -

DR EMBL; M60718; AAA52648.1; -

DR EMBL; X16323; CAA34387.1; -

DR EMBL; M73239; AAA64239.1; -

DR EMBL; M73240; AAA64297.1; -

DR EMBL; AC004960; AAC71655.1; -

DR EMBL; M75983; AAG53460.1; JOINED.

DR EMBL; M75972; AAG53460.1; JOINED.

DR EMBL; M75973; AAG53460.1; JOINED.

DR EMBL; M75974; AAG53460.1; JOINED.

DR EMBL; M75975; AAG53460.1; JOINED.

DR EMBL; M75976; AAG53460.1; JOINED.

DR EMBL; M75977; AAG53460.1; JOINED.

DR EMBL; M75978; AAG53460.1; JOINED.

DR EMBL; M75979; AAG53460.1; JOINED.

DR EMBL; M75980; AAG53460.1; JOINED.

DR EMBL; M75981; AAG53460.1; JOINED.

DR EMBL; M75982; AAG53460.1; JOINED.

DR PIR; JH0579; JH0579.

DR PIR; S06794; S06794.

DR PDB; 1BHT; 18-NOV-98.

DR MEROPS; S01.976; -

DR GlycoSuiteDB; P14210; -

DR Genew; HGNC:4893; HGF.

DR MIM; 142409; -

DR InterPro; IPR001314; Chymotrypsin.

DR InterPro; IPR000001; Kringle.

DR InterPro; IPR003014; PAN.

DR InterPro; IPR003609; Pan_app.

DR InterPro; IPR001254; Ser_protease_Try.

DR Pfam; PF00024; PAN; 1

DR Pfam; PF00051; kringle; 4.

DR Pfam; PF00089; trypsin; 1.

DR PRINTS; PR00722; CHYMOTRYPSIN.

DR PRINTS; PR00018; KRINGLE.

DR PRODOM; PD000395; Kringle; 4.

DR SMART; SM00130; KR; 4.

DR SMART; SM00473; PAN_AP; 1.

DR SMART; SM00020; TRY_PSP; 1.

DR PROSITE; PS00021; KRINGLE_1; 4.

DR PROSITE; PS00070; KRINGLE_2; 4.

DR PROSITE; PS02040; TRYPSIN_DOM; 1.

KW Growth factor; Kringle; Glycoprotein; Serine protease homolog;

KW Repeat; Signal; 3D-structure.

FT SIGNAL; 1 31

FT CHAIN; 32 494

FT CHAIN; 495 728

FT MOD_RES; 32 32

FT DOMAIN; 32 127

FT DOMAIN; 128 206

FT DOMAIN; 211 288

FT DOMAIN; 305 383

FT DOMAIN; 391 469

FT DOMAIN; 495 728

FT DISULFID; 70 96

FT DISULFID; 74 84

FT DISULFID; 128 206

FT DISULFID; 149 189

FT DISULFID; 177 201

HEPATOCYTE GROWTH FACTOR ALPHA CHAIN.

HEPATOCYTE GROWTH FACTOR BETA CHAIN.

PYRROLIDONE CARBOXYLIC ACID.

PAN.

KRINGLE 1.

KRINGLE 2.

KRINGLE 3.

KRINGLE 4.

SERINE PROTEASE-LIKE.

Query Match	99.9%; Score 2601; DB 1; Length 728;	
Best Local Similarity	99.8%; Pred. No. 8.3e-192;	
Matches 446; Conservative 1; Mismatches 0; Gaps 0;		
Oy 1	ERKRRNTTHERKSAKTLIKIDPALKTKTKVNTADOCANCRCTRNKGLPTCKAFVFDK 60	
Db 32	QKRRNTTHERKSAKTLIKIDPALKTKTKVNTADOCANCRCTRNKGLPTCKAFVFDK 91	
Oy 61	ARKQCLWFPFNSMSGVKKFEGHEFDLYENKDYIRNCIIIGKRSYKGTVSITKSGIKQCP 120	
Db 92	ARKQCLWFPFNSMSGVKKFEGHEFDLYENKDYIRNCIIIGKRSYKGTVSITKSGIKQCP 151	
Oy 121	WSSMIPHEHSELPSSYRGKDLQENYCRNPRGEEGPMWFTSNPEVRYEVCDDIPQCEVEEC 180	
Db 152	WSSMIPHEHSELPSSYRGKDLQENYCRNPRGEEGPMWFTSNPEVRYEVCDDIPQCEVEEC 211	
Oy 181	MTNCGESYRGLMDHTESKICQRDHQTTPHRRKFLPERYPDKGFDNCRNPDGQPRWC 240	
Db 212	MTNCGESYRGLMDHTESKICQRDHQTTPHRRKFLPERYPDKGFDNCRNPDGQPRWC 271	
Oy 241	YTLDPHTRWECYCAKTCADNTMNDTVDPLETECIQOGEGYRGVNTIWNIPQORWDS 300	
Db 272	YTLDPHTRWECYCAKTCADNTMNDTVDPLETECIQOGEGYRGVNTIWNIPQORWDS 331	
Oy 301	QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 360	
Db 332	QYPHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSHGQDC 391	
Oy 361	YRGNGKNTYMGNSQTSRGLTCSMDKXNEDLHRHFWPEPDASKLNENYCRNPDHAGPW 420	
Db 392	YRGNGKNTYMGNSQTSRGLTCSMDKXNEDLHRHFWPEPDASKLNENYCRNPDHAGPW 451	
Oy 421	CYTGNPLIPWDYCPISRCGEGDTTPIV 447	
Db 452	CYTGNPLIPWDYCPISRCGEGDTTPIV 478	
RESULT 2		
ID	HGF_MOUSE	STANDARD; PRP; 728 AA.
AC	O08048; O64007; O61662;	
DT	01-NOV-1995 (Rel. 32, Created)	
DT	01-NOV-1995 (Rel. 32, Last sequence update)	
DT	15-JUN-2002 (Rel. 41, Last annotation update)	
DE	Hepatocyte growth factor precursor (Scatter factor) (SF)	
DE	(Hepatopoietin A).	
GN	HGF.	
OS	Mus musculus (Mouse).	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	
OX	NCBI_TaxID=10090;	
RN	[1]	
RP	SEQUENCE FROM N.A. (LONG AND SHORT ISOFORMS), AND SEQUENCE OF 496-504.	
RC	TISSUE=Mammary fibroblast;	
RX	MEDLINE=94183257; PubMed=8135822;	
RA	Sasaki M., Nishio M., Sasaki T., Enami J.;	
RA	"Identification of mouse mammary fibroblast-derived mammary growth	
RT	factor as hepatocyte growth factor.";	
RL	Biochem. Biophys. Res. Commun. 199:772-779(1994).	
RN	[2]	
RP	SEQUENCE FROM N.A.	
RC	TISSUE=Liver;	
RA	MEDLINE=94363381; PubMed=8081873;	
RA	Lee C.C., Kozak C.A., Yamada K.M.;	
RT	"Structure, genetic mapping, and expression of the mouse Hgf/scatter	
RT	factor gene.";	
RL	Cell Adhes. Commun. 1:101-111(1993).	
RN	[3]	
RP	SEQUENCE FROM N.A.	
RC	TISSUE=Liver;	
RX	MEDLINE=94060105; PubMed=8241272;	
RA	Liu Y., Michalopoulos G.K., Zarnegar R.;	

Query Match

92.5%; Score 2408; DB 1; Length 728;

RT	"Molecular cloning and characterization of cDNA encoding mouse	
RL	hepatocyte growth factor.";	
CC	Biochim. Biophys. Acta 1216:299-303(1993).	
CC	-!- FUNCTION: HGF IS A POTENT MITOGEN FOR MATURE PARENCHYMAL	
CC	HEPATOCYTE CELLS, SEEMS TO BE AN HEPATOTROPIC FACTOR, AND ACTS	
CC	AS GROWTH FACTOR FOR A BROAD SPECTRUM OF TISSUES AND CELL TYPES.	
CC	IT HAS NO DETECTABLE PROTEASE ACTIVITY.	
CC	-!- SUBUNIT: DIMER OF AN ALPHA CHAIN AND A BETA CHAIN LINKED BY A	
CC	DISULFIDE BOND.	
CC	-!- ALTERNATIVE PRODUCTS: 2 isoforms; a long form (shown here) and a	
CC	short form; are produced by alternative splicing.	
CC	-!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. PLASMINOGEN SUBFAMILY.	
CC	-!- SIMILARITY: CONTAINS 4 KRINGLE DOMAINS.	
CC	-----	
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration	
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation -	
CC	the European Bioinformatics Institute. There are no restrictions on its	
CC	use by non-profit institutions as long as its content is in no way	
CC	modified and this statement is not removed. Usage by and for commercial	
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/	
CC	or send an email to license@isb-sib.ch).	
CC	-----	
DR	EMBL; D10212; BAA01064.1; -	
DR	EMBL; D10213; BAA01065.1; -	
DR	EMBL; S71816; AAB31855.1; -	
DR	EMBL; X72307; CAA51054.1; ALT_INIT.	
DR	HSSP; P14210; 1BHT.	
DR	MEROPS; S01.982; -	
DR	MGI; 96079; Hgf.	
DR	InterPro; IPR001314; Chymotrypsin.	
DR	InterPro; IPR000001; Kringle.	
DR	InterPro; IPR003014; PAN.	
DR	InterPro; IPR003609; Pan_app.	
DR	InterPro; IPR001254; Ser_protease_Try.	
DR	Pfam; PF00024; PAN. 1.	
DR	Pfam; PF00051; kringle; 4.	
DR	Pfam; PF00089; trypsin; 1.	
DR	PRINTS; PR00722; CHYMOTRYPSIN.	
DR	PRINTS; PR00018; KRINGLE.	
DR	ProDom; PD000395; Kringle; 4.	
DR	SMART; SM00130; KR; 4.	
DR	SMART; SM00473; PAN_AP; 1.	
DR	SMART; SM00020; TRY_SPC; 1.	
DR	PROSITE; PS00021; KRINGLE_1; 4.	
DR	PROSITE; PS00070; KRINGLE_2; 4.	
DR	PROSITE; PS0240; TRYPSIN_DOM; 1.	
KW	Growth factor; Kringle; Glycoprotein; Serine protease homolog;	
KW	Repeat; Signal; Alternative splicing.	
FT	SIGNAL 1 32	BY SIMILARITY.
FT	CHAIN 33 495	HEPATOCYTE GROWTH FACTOR ALPHA CHAIN.
FT	CHAIN 496 728	HEPATOCYTE GROWTH FACTOR BETA CHAIN.
FT	MOD_RES 33 33	PYROLIDONE CARBOXYLIC ACID
FT		(BY SIMILARITY).
FT	DOMAIN 33 128	PAP.
FT	DOMAIN 129 207	KRINGLE 1.
FT	DOMAIN 212 289	KRINGLE 2.
FT	DOMAIN 306 384	KRINGLE 3.
FT	DOMAIN 392 470	KRINGLE 4.
FT	DOMAIN 496 728	SERINE PROTEASE-LIKE.
FT	DISULFID 71 97	BY SIMILARITY.
FT	DISULFID 75 85	BY SIMILARITY.
FT	DISULFID 488 607	INTERCHAIN (BY SIMILARITY).
FT	CARBOHYD 295 295	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD 403 403	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD 569 569	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD 656 656	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	VARSPIC 163 167	MISSING (IN SHORT ISOFORM).
FT	CONFLICT 344 344	N -> K (IN REF. 2).
FT	CONFLICT 479 479	V -> L (IN REF. 2).
FT	CONFLICT 564 564	R -> H (IN REF. 3).
SQ	SEQUENCE 728 AA: 82944 MW; A0381FC497534328 CRC64;	

```
Best Local Similarity 91.1%, Pred. No. 4.9e-177;
Matches 407; Conservative 21; Mismatches 19; Indels 0; Gaps 0;

QY 1 EKKRRNTLHEFKSAKTLIKIDPAKTKTKVNTADOCANCRTRNKGLPFTCKAFVFDK 60
DB 33 OKKRRNTLHEFKSAKTLTKEDPLKTKTKVNSADCANCRIRNRGFTFTCKAFVFDK 92
QY 61 ARKQCLWPFNSMSGVKGKGEHFDLYENKDYIRNCIIIGKGRYKGVSTKSGIKQCP 120
DB 93 SRKRCYWPFPNSMSGVKGKGEHFDLYENKDYIRNCIIIGKGRYKGVSTKSGIKQCP 152
QY 121 WSMPTIPEHSHFLPSSYRGKDLQENYCRNPRGEGGPGWFTSNPEVRYEVCIDIPOCSEVEC 180
DB 153 WNSMPTIPEHSHFLPSSYRGKDLQENYCRNPRGEGGPGWFTSNPEVRYEVCIDIPOCSEVEC 212
QY 181 MTCNGESYRGLMDHTESGKICQWHDQTPHRRKFLPERYPDKGFDNDYCRNPDGQPRPWC 240
DB 213 MTCNGESYRGLMDHTESGKICQWHDQTPHRRKFLPERYPDKGFDNDYCRNPDGQPRPWC 272
QY 241 YTLDPHTRWEYCAIKTCADNTMDNDVPLETTECICQOGEGYRGTVNTIWNIGIPQWRDS 300
DB 273 YTLDPHTRWEYCAIKTCADNTMDNDVPLETTECICQOGEGYRGTVNTIWNIGIPQWRDS 332
QY 301 QYPHEHDMTPENFKCKDLRENYCRNPDSGSPWCTTDPNIRVGYCSQIPNCDMSHGQDC 360
DB 333 QYPHEHDMTPENFKCKDLRENYCRNPDSGSPWCTTDPNIRVGYCSQIPNCDMSHGQDC 392
QY 361 YRCNGKNYGNLSQPSGLTCSMDKNDLHRHFWPEPDASKLKNENYCRNPDDDAHGFW 420
DB 393 YRCNGKNYGNLSQPSGLTCSMDKNDLHRHFWPEPDASKLKNENYCRNPDDDAHGFW 452
QY 421 CYTGNPLIPWDYCPISRCGSDTTPIV 447
DB 453 CYTGNPLIPWDYCPISRCGSDTTPIV 479

RESULT 3
ID HGF_RAT STANDARD; PRT; 728 AA.
AC P17945;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hepatocyte growth factor precursor (Scatter factor) (SF)
DE (Hepatopoietin A).
GN HGF
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
ON NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC STRAIN=Wistar; TISSUE=Liver;
RX MEDLINE=90222197; PubMed=2139229;
RA Toshiro K., Hagiya M., Nishizawa T., Seki T., Shimonishi M.,
RA Shimizu S., Nakamura T.;
RT "deduced primary structure of rat hepatocyte growth factor and
RT expression of the mRNA in rat tissues.";
RL Proc. Natl. Acad. Sci. U.S.A. 87:3200-3204(1990).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Wistar; TISSUE=Liver;
RX MEDLINE=91031482; PubMed=2146117;
RA Okajima A., Miyazawa K., Kitamura N.;
RT "primary structure of rat hepatocyte growth factor and induction of
RT its mRNA during liver regeneration following hepatic injury.";
RL Eur. J. Biochem. 193:375-381(1990).
CC -1- FUNCTION: HGF IS A POTENT MITOGEN FOR MATURE PARENCHYMAL
CC HEPATOCYTE CELLS, SEEMS TO BE AN HEPATOTROPIC FACTOR, AND ACTS
CC AS GROWTH FACTOR FOR A BROAD SPECTRUM OF TISSUES AND CELL TYPES.
CC IT HAS NO DETECTABLE PROTEASE ACTIVITY.
CC -1- SUBUNIT: DIMER OF AN ALPHA CHAIN AND A BETA CHAIN LINKED BY A
CC DISULFIDE BOND.
```

```
CC CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. PLASMINOGEN SUBFAMILY.
CC CC -1- SIMILARITY: CONTAINS 4 KRINGLE DOMAINS.
CC CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC CC or send an email to license@isb-sib.ch).
CC CC -----
CC CC EMBL; D90102; BAA14133.1;
CC CC EMBL; X54400; CAA38266.1;
CC CC PIR; S13211; S13211.
CC CC PIR; A35644; A35644.
CC CC HSP; P14210; IAH.
CC CC MEROPS; S01.978;
CC CC InterPro; IPR001314; Chymotrypsin.
CC CC InterPro; IPR000001; Kringle.
CC CC InterPro; IPR003014; PAN.
CC CC InterPro; IPR003609; Pan_app.
CC CC InterPro; IPR001254; Ser_protease_Try.
CC CC Pfam; PF00024; PAN; 1.
CC CC Pfam; PF00051; kringle; 4.
CC CC Pfam; PF00089; trypsin; 1.
CC CC PRINTS; PR00722; CHYMOTRYPSIN.
CC CC PRINTS; PR00018; KRINGLE.
CC CC ProDom; PD000395; Kringle; 4.
CC CC SMART; SM00130; KR; 4.
CC CC SMART; SM00473; PAN_AP; 1.
CC CC SMART; SM00020; Tryp_Spc; 1.
CC CC PROSITE; PS00021; KRINGLE_1; 4.
CC CC PROSITE; PS00070; KRINGLE_2; 4.
CC CC PROSITE; PS0240; TRYPSIN_DOM; 1.
CC CC Growth factor; Kringle; Glycoprotein; Serine protease homolog;
CC KW Repeat; Signal; 1.
CC FT SIGNAL 1 32 BY SIMILARITY.
CC FT CHAIN 33 495 HEPATOCYTE GROWTH FACTOR ALPHA CHAIN.
CC FT CHAIN 496 728 HEPATOCYTE GROWTH FACTOR BETA CHAIN.
CC FT MOD_RES 33 33 PYROLIDONE CARBOXYLIC ACID
CC FT (BY SIMILARITY).
CC FT DOMAIN 33 128 PAP.
CC FT DOMAIN 129 207 KRINGLE 1.
CC FT DOMAIN 212 289 KRINGLE 2.
CC FT DOMAIN 306 384 KRINGLE 3.
CC FT DOMAIN 392 470 KRINGLE 4.
CC FT DOMAIN 496 728 SERINE PROTEASE-LIKE.
CC FT DISULFID 71 97 BY SIMILARITY.
CC FT DISULFID 75 85 BY SIMILARITY.
CC FT DISULFID 488 607 INTERCHAIN (BY SIMILARITY).
CC FT CARBOHYD 295 295 N-LINKED (GLCNAC. .) (POTENTIAL).
CC FT CARBOHYD 403 403 N-LINKED (GLCNAC. .) (POTENTIAL).
CC FT CARBOHYD 569 569 N-LINKED (GLCNAC. .) (POTENTIAL).
CC FT CARBOHYD 656 656 N-LINKED (GLCNAC. .) (POTENTIAL).
CC SQ SEQUENCE 728 AA; 82905 MW; 350BF1F96ADCDEF CRC64;
Query Match 92.2%; Score 2401; DB 1; Length 728;
Best Local Similarity 90.4%; Pred. No. 1.7e-176;
Matches 404; Conservative 23; Mismatches 20; Indels 0; Gaps 0;
QY 1 EKKRRNTLHEFKSAKTLIKIDPAKTKTKVNTADOCANCRTRNKGLPFTCKAFVFDK 60
DB 33 OKKRRNTLHEFKSAKTLTKEDPLKTKTKVNSADCANCRIRNRGFTFTCKAFVFDK 92
QY 61 ARKQCLWPFNSMSGVKGKGEHFDLYENKDYIRNCIIIGKGRYKGVSTKSGIKQCP 120
DB 93 SRKRCYWPFPNSMSGVKGKGEHFDLYENKDYIRNCIIIGKGRYKGVSTKSGIKQCP 152
QY 121 WSMPTIPEHSHFLPSSYRGKDLQENYCRNPRGEGGPGWFTSNPEVRYEVCIDIPOCSEVEC 180
DB 153 WNSMPTIPEHSHFLPSSYRGKDLQENYCRNPRGEGGPGWFTSNPEVRYEVCIDIPOCSEVEC 212
QY 181 MTCNGESYRGLMDHTESGKICQWHDQTPHRRKFLPERYPDKGFDNDYCRNPDGQPRPWC 240
```

Db 213 MTCNGESYRGPMDHTEGSKTCQRDQQTPIRHKFLPERYPKGFDDNYCRNPDGKPRPW 272
QY 241 YTLDPHPRWYCAIKTCADNTMDTDPLETTICOGEGYGTNTWNGIPICQWRDWS 300
Db 273 YTLDPDPWEYCAIKWCAHSAVNETDVPMETTECIKOGEGYGTNTWNGIPICQWRDWS 332
QY 301 QYEHEDMTPEFKCKDLRENYCRNPDGSSPMCFPTDPNIRVGYCSQIPNCDMSHGQDC 360
Db 333 QYPHKHDTPEFKCKDLRENYCRNPDGSAESPMCFPTDPNIRVGYCSQIPKCDVSSQDC 392
QY 361 YRGNKKNYGNLSQTRSGLTCSMWKKNMEDLHRHIFPEPDASKNLNENYCRNPDHAGPW 420
Db 393 YRGNKKNYGNLSKTRSGLTCSMWKKNMEDLHRHIFPEPDASKLTKNYCRNPDHAGPW 452
QY 421 CYTGNPLIPWDYCPISRCCEGDTPTIV 447
Db 453 CYTGNPLIPWDYCPISRCCEGDTPTIV 479
RESULT 4
HGFL_HUMAN STANDARD; PRT; 711 AA.
AC P26927: Q13350; Q14870;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hepatocyte growth factor-like protein precursor (Macrophage
stimulatory protein) (MSP) (Macrophage stimulating protein).
GN MST1 OR HGFL.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=92002016; PubMed=1655021;
RA Han S., Stuart L.A., Friezneger Degen S.J.;
RT "Characterization of the DNFI52 locus on human chromosome 3:
identification of a gene coding for four kringle domains with
homology to hepatocyte growth factor.";
RL Biochemistry 30:9768-9780(1991).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=93340141; PubMed=8393443;
RA Yoshimura T., Yuhki N., Wang M.H., Skeel A., Leonard E.J.;
RT "Cloning, sequencing, and expression of human macrophage stimulating
protein (MSP, MST1) confirms MSP as a member of the family of kringle
proteins and locates the MSP gene on chromosome 3.";
RL J. Biol. Chem. 268:15461-15468(1993).
CC -1- FUNCTION: PROBABLY HAS NO PROTEOLYTIC ACTIVITY, SINCE CRUCIAL AA
CHARACTERISTIC OF SERINE PROTEASES CATALYTIC SITES ARE NOT
CONSERVED.
CC -1- PTM: MAY BE CLEAVED AFTER AA 484, TO YIELD A TWO-CHAIN MOLECULE
HELLED TOGETHER BY DISULFIDE BONDS, OR TWO SEPARATE POLYPEPTIDES.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. PLASMINOGEN SUBFAMILY.
CC -1- SIMILARITY: CONTAINS 4 KRINGLE DOMAINS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M74178; AAA50165.1; -
DR EMBL; U37055; AAC50471.1; -
DR EMBL; L11924; AAA59872.1; -
DR PIR; A40331; A40331.
DR HSP; P00747; 2PK4.

DR MEROPS; S01.975; -
DR Genem; HGNC:7380; MST1.
DR MIM; 142408;
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR003014; PAN.
DR InterPro; IPR003609; Pan_app.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00024; PAN; 1.
DR Pfam; PF00051; kringle; 4.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR PRODOM; PD000395; Kringle; 4.
DR SMART; SM00130; KR; 4.
DR SMART; SM00473; PAN_AP; 1.
DR SMART; SM00020; TRYD_SPC; 1.
DR PROSITE; PS00021; KRINGLE_1; 4.
DR PROSITE; PS00070; KRINGLE_2; 4.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
KW Kringle; Glycoprotein; Serine protease homolog; Repeat; Signal;
KW Polymorphism.
FT SIGNAL 1 18 POTENTIAL.
FT CHAIN 19 711 HEPATOCYTE GROWTH FACTOR-LIKE PROTEIN.
FT DOMAIN 32 109 PAP.
FT DOMAIN 110 186 KRINGLE 1.
FT DOMAIN 191 268 KRINGLE 2.
FT DOMAIN 283 361 KRINGLE 3.
FT DOMAIN 370 448 KRINGLE 4.
FT DOMAIN 484 711 SERINE PROTEASE-LIKE.
FT DISULFID 56 78 BY SIMILARITY.
FT DISULFID 60 66 BY SIMILARITY.
FT DISULFID 110 186 BY SIMILARITY.
FT DISULFID 131 169 BY SIMILARITY.
FT DISULFID 157 181 BY SIMILARITY.
FT DISULFID 191 268 BY SIMILARITY.
FT DISULFID 194 324 INTERCHAIN (BY SIMILARITY).
FT DISULFID 212 251 BY SIMILARITY.
FT DISULFID 240 263 BY SIMILARITY.
FT DISULFID 283 361 BY SIMILARITY.
FT DISULFID 304 343 BY SIMILARITY.
FT DISULFID 332 355 BY SIMILARITY.
FT DISULFID 370 448 BY SIMILARITY.
FT DISULFID 391 431 BY SIMILARITY.
FT DISULFID 419 443 INTERCHAIN (BY SIMILARITY).
FT DISULFID 468 588 BY SIMILARITY.
FT DISULFID 507 523 BY SIMILARITY.
FT DISULFID 602 667 BY SIMILARITY.
FT DISULFID 632 646 BY SIMILARITY.
FT DISULFID 657 685 BY SIMILARITY.
FT CARBOHYD 72 72 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 296 296 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 615 615 N-LINKED (GLCNAC. .) (POTENTIAL).
FT VARIANT 13 13 Y -> C.
FT VARIANT 212 212 /FTID=VAR_006631.
FT VARIANT 623 623 C -> F.
FT CONFLICT 623 623 /FTID=VAR_006632.
FT CONFLICT 711 AA; 80379 MW; 596ED21F180290E4 CRC64;
SQ SEQUENCE 711 AA; 80379 MW; 596ED21F180290E4 CRC64;
Query Match 45.3%; Score 1180.5; DB 1; Length 711;
Best Local Similarity 47.4%; Pred. No. 4.4e-83;
Matches 197; Conservative 62; Mismatches 148; Indels 9; Gaps 4;
QY 33 VNTADQCANCRTRNKGLPETCKAFVFDKARKQCLWFFNSMSGVKKKEFGHEFDLYENKD 92
Db 50 VADDEECAGRC-----GLPMDCRAFHYNVSSHGCOLLPWTQHPHTRLRSGRCDFQKDD 105
QY 93 YIRNCIIIGKGRYKGTIVTSKGIKQCPWSSMIPHEHSFLPSSYRGKDLQENYCRNPRGE 152
Db 106 YVRTCINNGVGYRGTMATVGGPLPCQAWSHKFFNDHKYTTPLNG--LEENFCRNPDGD 163
QY 153 EGGPWCTSNPEVRYEVCDDIPQCSEVEMTCNGESYRGKLDHTESGKICQRDHQTPIRH 212

```
Db 164 PGGWCYTDDPAVFQSCGKSCREACWCNGEYGAVDRTESGECORWDLQHPQH 223
Qy 213 KFLPERYPDKGFDNYCRNPDGQPRPCWYTLDPHTRWEYCAIKTCADTNDWTDVPLETT 272
Db 224 PFECKFLDQGLDNYCRNPDGSRPCWYTTDPOIEREFCDLPRGSEAPQGEA--TTV 281
Qy 273 ECIQOGEGYGTVTIWNIGTPCORWDSQYPHEHDMTPENFKCKDLRENCRNPDGSESP 332
Db 282 SCFRKGEGYGTANTTAGVPCQWRDAQIPHQHRTPEKYACKDLRENFCRNPDGSEAP 341
Qy 333 WCFTDPNIRVGYQSQPN-C-DMSHGQDCYRGNGKNTMGNLSQTRSLGTCMMWKNMEDL 391
Db 342 WCFTLPCRMRAFCYQIRRCTDDVRPQDCYHAGEQYRGTVSKTRKGVCQQRWSAETPHK 401
Qy 392 HRHIFWEPDASKLENYCRNPDGDAHGPWCYTGPNLIPWDCYPSRCEGDTPTIV 447
Db 402 PQFTTSPHAQLEENFCRNPDSHGPGWCYTMDPRTPFYCALRRCAADDQPPSIL 457

RESULT 5
HGFL_MOUSE STANDARD; PRT; 716 AA.
AC P26928;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hepatocyte growth factor-like protein precursor (Macrophage
DE stimulatory protein) (MSP).
GN MST1 OR HGFL.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/C; TISSUE=Liver;
RX MEDLINE=92002017; PubMed=1832957;
RA Friezen-Degen S.J., Stuart L.A., Han S., Jamison C.S.;
RT "Characterization of the mouse cDNA and gene coding for a hepatocyte
RT growth factor-like protein: expression during development.";
RL Biochemistry 30:9781-9791(1991).
CC -!- FUNCTION: PROBABLY HAS NO PROTEOLYTIC ACTIVITY, SINCE CRUCIAL AA
CC CHARACTERISTIC OF SERINE PROTEASES CATALYTIC SITES ARE NOT
CC CONSERVED.
CC -!- TISSUE SPECIFICITY: LIVER. LOWER LEVELS IN LUNG, PLACENTA AND
CC ADRENAL.
CC -!- DEVELOPMENTAL STAGE: IS EXPRESSED AT LOW LEVELS DURING GESTATION.
CC STABLE BEFORE BIRTH THE LEVEL INCREASES DRAMATICALLY AND REMAINS
CC STABLE AFTERWARDS.
CC -!- PTM: MAY BE CLEAVED AFTER AA 488, TO YIELD A TWO-CHAIN MOLECULE
CC HELD TOGETHER BY DISULFIDE BONDS, OR CLEAVED INTO TWO SEPARATE
CC POLYPEPTIDES.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. PLASMINOGEN SUBFAMILY.
CC -!- SIMILARITY: CONTAINS 4 KRINGLE DOMAINS.

-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; M74180; AAA50166.1; -
EMBL; M74181; AAA50167.1; -
HSP; P00747; IKRN.
MEROPS; S01.975; -
MGD; MGI:96080; Hgfl.
InterPro; IPR001314; Chymotrypsin.
InterPro; IPR003001; Kringle.
InterPro; IPR003014; PAN.
InterPro; IPR003609; Pan_app.
```

```
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00024; PAN; 1.
DR Pfam; PF00051; kringle; 4.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PRO0722; CHYMOTRYPSIN.
DR PRINTS; PRO0018; KRINGLE.
DR ProDom; PD000395; Kringle; 4.
DR SMART; SM00130; KR; 4.
DR SMART; SM00473; PANAP; 1.
DR SMART; SM00020; TRYD-SPC; 1.
DR PROSITE; PS00021; KRINGLE_1; 4.
DR PROSITE; PS0070; KRINGLE_2; 4.
DR PROSITE; PS0240; TRYPSIN_DOM; 1.
KW Kringle; Glycoprotein; Serine protease homolog; Repeat; Signal.
FT SIGNAL 1 18
FT CHAIN 19 716
FT DOMAIN 19 109
FT DOMAIN 110 186
FT DOMAIN 191 268
FT DOMAIN 292 370
FT DOMAIN 379 457
FT DOMAIN 489 716
FT DISULFID 56 78
FT DISULFID 60 66
FT DISULFID 110 186
FT DISULFID 131 169
FT DISULFID 157 181
FT DISULFID 191 268
FT DISULFID 194 333
FT DISULFID 212 251
FT DISULFID 240 263
FT DISULFID 292 370
FT DISULFID 313 352
FT DISULFID 341 364
FT DISULFID 379 457
FT DISULFID 400 440
FT DISULFID 428 452
FT DISULFID 477 593
FT DISULFID 512 528
FT DISULFID 607 672
FT DISULFID 637 651
FT DISULFID 662 690
FT CARBOHYD 72 72
FT CARBOHYD 173 173
FT CARBOHYD 305 305
FT CARBOHYD 620 620
FT CONFLICT 19 19
SQ SEQUENCE 716 AA; 80588 MW; BBCE02EF85213ACC CRC64;

Query Match 43.28; Score 1125; DB 1; Length 716;
Best Local Similarity 45.98; Pred. No. 7.9e-79;
Matches 195; Conservative 69; Mismatches 143; Indels 18; Gaps 6;

Qy 33 VNTADOCANRCTRNKGLPFTCKAFVDFKARKQCLWFFNFSMSSGVKKKEFGHEFDLYENKD 92
Db 50 VADAEACARR-----GPLLDCRAFYHNNSSHCQQLLPWQHSHTQLYHSSLCHFQKDK 105
Qy 93 YIRNCIIGKRSYGTYSITKSGIKQCPWSSMIFHEHSFLPSSYRGKDLQNYCRNPRGE 152
Db 106 YVRTCIMDNGVSYRGTVARTAGGLPCQAWSRRFFNDHKYTPTPKNG--LEENFCRNPDDG 163
Qy 153 EGGPWCFTSNVEVRYEVCDDIPQCSEVCMTCNGESYRGLMDHTESGKTCORWDHOTPHR 212
Db 164 PRGPWCYTNNRSVFSQSGIKTCREAVCVLCNGEDYRGVDVTEGREGCQWDLQHPHSH 223
Qy 213 KFLPERYPDKGFDNYCRNPDGQPRPCWYTLDPHTRWEYCAIKTCADN-----TMDT 265
Db 224 PQPEKFLDKDLKDYCRNPDGSRPCWYTTDPNVEREFCDLPCGPNLPPTVKGSKSQ 283
Qy 266 DVPLETTECIQOGEGYGTVTIWNIGTPCORWDSQYPHEHDMTPENFKCKDLRENCRN 325
Db 284 RNKGKALNCFRGKGDYRGTTNTTSAGVPCQWRDAQSPHQHRTPEKYACKDLRENFCRN 343
```


QY 422 YTCNPLIPWDYCPISRC 438
ID PLMN_MOUSE STANDARD; PRT; 812 AA.
AC P20918;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Plasminogen precursor (EC 3.4.21.7) [Contains: Angiostatin].
GN PLG.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1] SEQUENCE FROM N.A.
RP MEDLINE=91184812; PubMed=2081600;
RX MEDLINE=95042728; PubMed=7525077;
RA O'Reilly M.S., Holmgren L., Shing Y., Chen C., Rosenthal R.A.,
RA Moses M., Lane W.S., Cao Y., Sage E.H., Folkman J.;
RT "Angiostatin: a novel angiogenesis inhibitor that mediates the
RT suppression of metastases by a Lewis lung carcinoma."
RL Cell 79:315-328(1994).
CC -!- FUNCTION: PLASMIN DISSOLVES THE FIBRIN OF BLOOD CLOTS AND ACTS AS
CC A PROTEOLYTIC FACTOR IN A VARIETY OF OTHER PROCESSES INCLUDING
CC EMBRYONIC DEVELOPMENT, TISSUE REMODELING, TUMOR INVASION,
CC AND INFLAMMATION; IN OVULATION IT WEAKENS THE WALLS OF THE
CC GRAAFIAN FOLLICLE. IT ACTIVATES THE UROKINASE-TYPE PLASMINOGEN
CC ACTIVATOR, COLLAGENASES AND SEVERAL COMPLEMENT ZYMOGENS, SUCH
CC AS C1 AND C5. IT CLEAVES FIBRIN, FIBRONECTIN, THROMBOSPONDIN,
CC LAMININ AND VON WILLEBRAND FACTOR.
CC -!- FUNCTION: ANGIOSTATIN IS AN ANGIOGENESIS INHIBITOR THAT BLOCKS
CC NEOVASCULARIZATION AND GROWTH OF EXPERIMENTAL PRIMARY AND
CC METASTATIC TUMORS IN VIVO.
CC -!- CATALYTIC ACTIVITY: Preferential cleavage: Lys-I-Xaa > Arg-I-Xaa;
CC higher selectivity than trypsin. Converts fibrin into soluble
CC products.
CC -!- ENZYME REGULATION: CONVERTED INTO PLASMIN BY PLASMINOGEN
CC ACTIVATORS, BOTH PLASMINOGEN AND ITS ACTIVATOR BEING BOUND TO
CC FIBRIN. CANNOT BE ACTIVATED WITH STREPTOKINASE.
CC -!- MISCELLANEOUS: PLASMIN IS INACTIVATED BY ALPHA-2-ANTIPLASMIN
CC IMMEDIATELY AFTER DISSOCIATION FROM THE CLOT.
CC -!- MISCELLANEOUS: IN THE PRESENCE OF THE INHIBITOR, THE ACTIVATION
CC INVOLVES ONLY CLEAVAGE AFTER ARG-581, RESULTING IN 2 CHAINS HELD
CC TOGETHER BY 2 DISULFIDE BONDS. WITHOUT THE INHIBITOR, THE
CC ACTIVATION INVOLVES ALSO REMOVAL OF THE ACTIVATION PEPTIDE.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. PLASMINOGEN SUBFAMILY.
CC -!- SIMILARITY: CONTAINS 5 KRINGLE DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; J04766; AAA50168.1; -
CC PIR; A38514; A38514.
CC HSSP; P00747; 1PMK.
CC MEROPS; S01.233; -
CC MGD; MGI:97620; Plg.
CC InterPro; IPR001314; Chymotrypsin.

DR InterPro: IPR000001; Kringle.
DR InterPro: IPR003014; PAN.
DR InterPro: IPR003609; Pan_app.
DR InterPro: IPR001254; Ser_protease_Try.
DR Pfam; PF00024; PAN; 1.
DR Pfam; PF00051; kringle; 5.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PRO0722; CHYMOTRYPSIN.
DR PRINTS; PRO0018; KRINGLE.
DR ProDom; PD000395; Kringle; 5.
DR SMART; SM00130; KR; 5.
DR SMART; SM00473; PAN_AP; 1.
DR SMART; SM00420; TRYPSIN; 1.
DR PROSITE; PS00021; KRINGLE_1; 4.
DR PROSITE; PS00070; KRINGLE_2; 5.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Serine protease; Plasma; Glycoprotein; Fibrinolysis;
KW Tissue remodeling; Blood coagulation; Kringle; Zymogen; Repeat;
KW Signal.
FT SIGNAL 1 19 PLASMINOGEN.
FT CHAIN 20 812 PLASMIN HEAVY CHAIN A.
FT CHAIN 20 581 ACTIVATION PEPTIDE.
FT PEPTIDE 20 97 PLASMIN SHORT FORM OF CHAIN A.
FT CHAIN 98 581 PLASMINOGEN.
FT CHAIN 98 7436 ANGIOSTATIN.
FT CHAIN 582 812 PLASMIN LIGHT CHAIN B.
FT DOMAIN 103 181 KRINGLE 1.
FT DOMAIN 184 262 KRINGLE 2.
FT DOMAIN 275 352 KRINGLE 3.
FT DOMAIN 377 454 KRINGLE 4.
FT DOMAIN 481 560 KRINGLE 5.
FT DOMAIN 582 812 SERINE PROTEASE.
FT ACT_SITE 624 624 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 667 667 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 762 762 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT DISULFID 49 73 BY SIMILARITY.
FT DISULFID 53 61 BY SIMILARITY.
FT DISULFID 103 181 BY SIMILARITY.
FT DISULFID 124 164 BY SIMILARITY.
FT DISULFID 152 176 BY SIMILARITY.
FT DISULFID 185 262 BY SIMILARITY.
FT DISULFID 208 316 BY SIMILARITY.
FT DISULFID 206 245 BY SIMILARITY.
FT DISULFID 234 257 BY SIMILARITY.
FT DISULFID 275 335 BY SIMILARITY.
FT DISULFID 296 335 BY SIMILARITY.
FT DISULFID 324 347 BY SIMILARITY.
FT DISULFID 377 454 BY SIMILARITY.
FT DISULFID 398 437 BY SIMILARITY.
FT DISULFID 426 449 BY SIMILARITY.
FT DISULFID 481 560 BY SIMILARITY.
FT DISULFID 502 543 BY SIMILARITY.
FT DISULFID 531 555 BY SIMILARITY.
FT DISULFID 568 687 INTERCHAIN (BY SIMILARITY).
FT DISULFID 578 586 INTERCHAIN (BY SIMILARITY).
FT DISULFID 609 625 BY SIMILARITY.
FT DISULFID 701 768 BY SIMILARITY.
FT DISULFID 731 747 BY SIMILARITY.
FT DISULFID 758 786 BY SIMILARITY.
SQ SEQUENCE 812 AA; 90846 MW; D34A74A4FC2256F8 CRC64;
Query Match 40.7%; Score 1059; DB 1; Length 812;
Best Local Similarity 43.3%; Pred.No. 1e-73;
Matches 191; Conservative 67; Mismatches 145; Indels 38; Gaps 12;
QY 30 TKKVNTA---DQCANRCTRNGKLPFTCKAFVDFKARKQCLWPFNMSGVKKEFGHEFD 86
DB 37 TKKLAAGVSDCLAKEGE--TDFVCRSFQVHSKEQOCVMAENSKTSSIRN--RDVI 92
QY 87 LYENKDYIRNCIIGKGRSYKGVSTKSIKQCPWSSMIPHEHSLPSSYRGKDLQENVC 146

Db 93 LFEKRVYLSECKTGIGNCYRGTMTGKSGVACQKAGTFFHVPVNPSTHNEGLENYC 152

Qy 147 RNPGEEGPWCFTSNPEVRYEVCDDIPQCSVEVECMTCNGESYRGLMDHTSGKICQWRDH 206

Db 153 RNPNDDEGPCWYTTDPKDYDCNIPCEE-ECMYCSGEYEGKISKITMSGLDCQAWDS 211

Qy 207 QTPRHKFLPERYPDGKFDNDYCRNPDGQPRWCYTLDPHTRWEYCAIKTCAADTMNDT 266

Db 212 QSPHAGYIPAKFTSKNLKMYCNHPDGEPRWCFTTDPTRWEYCDIPRCT-TPPP 267

Qy 267 VPLETTCIOGEGYRGYRTVNTIWNIPQCRWDQYRPHEDMTPEFCKDLRENYCRNP 326

Db 268 PPSPTYQLKRGENYRGTSTVSGTKQCRWSEOTPHRHRTPEFCKDLRENYCRNP 327

Qy 327 DGSESPWCFTTDPNIRVGYCQIPNCDMSHG-----QDCYRGNKKNYM 369

Db 328 DGEAPWCYTTDSLRWEYC-EIPSCSSASPQSDSDSVPEEQTPVVQECYQSDGQSYR 386

Qy 370 GNLQTSRGLTCSMKWKNMEDLHRHIFWE---PDASKLNEYCRNPDDDAHPWCYTCNP 426

Db 387 GTSSTTIPTGKRCQSWAAMP--HRHSKTPENFPDAG-LEMNYCRNPDDG-KGPWCYTTDP 442

Qy 427 LIPWDYCPISRCGDTTPTIV 447

Db 443 SVRWEYCNLKKC-SETGGSV 462

RESULT 8

PLMN_ERIEU STANDARD; PRF: 810 AA.

AC Q29483;

DT 01-NOV-1997 (Rel. 35, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Plasminogen precursor (EC 3.4.21.7).

GN PLG.

OS Erinnaceus europaeus (Western European hedgehog).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Insectivora; Erinaceidae; Erinaceinae; Erinaceus.

OX NCBI_TaxID=9365;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Liver;

RA MEDLINE=96025778; PubMed=7592597;

RA Lawn R.M., Boommark N.W., Schwartz K., Lindahl G.E., Wade D.P.,

RA Byrne C.D., Fong K.J., Meer K., Patthy L.;

RT "The recurring evolution of lipoprotein(a)". Insights from cloning of

RT hedgehog apolipoprotein(a)".

RL J. Biol. Chem. 270:24004-24009(1995).

RN [2]

RP REVISIONS.

RA Lawn R.M.;

RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.

CC -1- FUNCTION: PLASMIN DISSOLVES THE FIBRIN OF BLOOD CLOTS AND ACTS AS

CC A PROTEOLYTIC FACTOR IN A VARIETY OF OTHER PROCESSES INCLUDING

CC EMBRYONIC DEVELOPMENT, TISSUE REMODELING, TUMOR INVASION,

CC AND INFLAMMATION; IN OVULATION IT WEAKENS THE WALLS OF THE

CC GRAAFIAN FOLLICLE. IT ACTIVATES THE UROKINASE-TYPE PLASMINOGEN

CC ACTIVATOR, COLLAGENASES AND SEVERAL COMPLEMENT ZYMOGENS. SUCH

CC AS C1 AND C5. IT CLEAVES FIBRIN, FIBRONECTIN, THROMBOSPONDIN,

CC LAMININ AND VON WILLEBRAND FACTOR.

CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Lys-|-Xaa > Arg-|-Xaa;

CC higher selectivity than trypsin. Converts fibrin into soluble

CC products.

CC -1- ENZYME REGULATION: CONVERTED INTO PLASMIN BY PLASMINOGEN

CC ACTIVATORS, BOTH PLASMINOGEN AND ITS ACTIVATOR BEING BOUND TO

CC FIBRIN. CANNOT BE ACTIVATED WITH STREPTOKINASE.

CC -1- MISCELLANEOUS: PLASMIN IS INACTIVATED BY ALPHA-2-ANTIPLASMIN

CC IMMEDIATELY AFTER DISSOCIATION FROM THE CLOT.

CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. PLASMINOGEN SUBFAMILY.

CC -1- SIMILARITY: CONTAINS 5 KRINGLE DOMAINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL; U33171; AAC48717.1;

DR HESP; P00747; LPMK.

DR MEROPS; S01.233;

DR InterPro; IPR001314; Chymotrypsin.

DR InterPro; IPR000001; Kringle.

DR InterPro; IPR003014; PAN.

DR InterPro; IPR003609; Pan_app.

DR InterPro; IPR001254; Ser_protease_Try.

DR Pfam; PF00024; PAN; 1.

DR Pfam; PF00051; kringle; 5.

DR Pfam; PF00089; trypsin; 1.

DR PRINTS; PR00722; CHYMOTRYPSIN.

DR PRINTS; PR00018; KRINGLE.

DR PRODOM; PD000395; Kringle; 5.

DR SMART; SM00130; KR; 5.

DR SMART; SM00473; PAN_AP; 1.

DR SMART; SM00020; Tryp_Spc; 1.

DR PROSITE; PS00021; KRINGLE_1; 5.

DR PROSITE; PS00070; KRINGLE_2; 5.

DR PROSITE; PS02040; TRYPsin_DOM; 1.

DR PROSITE; PS00134; TRYPsin_HIS; 1.

DR PROSITE; PS00135; TRYPsin_SER; 1.

DR Hydrolase; Serine protease; Plasma; Glycoprotein; Fibrinolysis;

KW Tissue remodeling; Blood coagulation; Kringle; Zymogen; Repeat;

KW Signal.

FT CHAIN 1 19 BY SIMILARITY.

FT CHAIN 20 810 PLASMINOGEN.

FT CHAIN 20 582 PLASMIN HEAVY CHAIN A (BY SIMILARITY).

FT CHAIN 583 810 PLASMIN LIGHT CHAIN B (BY SIMILARITY).

FT DOMAIN 583 810 SERINE PROTEASE.

FT DOMAIN 103 181 KRINGLE 1.

FT DOMAIN 185 262 KRINGLE 2.

FT DOMAIN 275 352 KRINGLE 3.

FT DOMAIN 379 456 KRINGLE 4.

FT DOMAIN 482 561 KRINGLE 5.

FT ACT_SITE 622 622 CHARGE RELAY SYSTEM.

FT ACT_SITE 665 665 CHARGE RELAY SYSTEM.

FT ACT_SITE 760 760 CHARGE RELAY SYSTEM.

FT CARBOHYD 339 339 N-LINKED (GLCNAC...) (POTENTIAL).

SQ SEQUENCE 810 AA; 90902 MW; 8E75780946017A16 CRC64;

Query Match 40.5%; Score 1053.5; DB 1; Length 810;

Best Local Similarity 44.3%; Pred. No. 2.7e-73;

Matches 194; Conservative 63; Mismatches 138; Indels 43; Gaps 13;

Qy 30 TKK---VNTADOCANRCTRNKGLPTCKAFVFDKARKOCLMFPFNSMSSGVKKFGEHFD 86

Db 37 TKKOLSVGTEECACVK--EKETSFICRSFYHSKQCVMAENSKSTPVLRM--RDVI 92

Qy 87 LYENKDYIRNCIIIGKRSYKGTIVITKSGIKQCPWSSMTPEHSPSPSYRKDLQENVC 146

Db 93 LFEKMYLSECKVNGKYGRTVSKTKTLGTCOKWSAETPHKPRESPDENEGLDQNYC 152

Qy 147 RNPGEEGPWCFTSNPEVRYEVCDDIPQCSVEVECMTCNGESYRGLMDHTSGKICQWRDH 206

Db 153 RNPNDDEGPCWYTTDPKDYDCNIPCEE-EDECCHCSQNYVGVKISRTMSGLEQCPWDS 211

Qy 207 QTPRHKFLPERYPDGKFDNDYCRNPDGQPRWCYTLDPHTRWEYCAIKTCAADTMNDT 266

Db 212 QIPHPGFIPIPSKFPKSNLKMNYCRNPDPGEPRWCFTMORNKRWEYCDIPRCT-TP 265

Qy 267 VPLE--TTECTGQGEYRGYRTVNTIWNIPQCRWDQYRPHEDMTPEFCKDLRENYCR 324

Db 266 PPSPTYQLKMGNGEYHOGVAVNTVSGLTQCRWGEQSPHRDTPENYPCKNLDENYCR 325

Qy 325 NPDGSESPWCFTTDPNIRVGYCQIPNC-----DMSGH-----QDCYRNG 365

RT "Evidence for a novel O-linked sialylated trisaccharide on Ser-248 of
human plasminogen 2.";
J. Biol. Chem. 272:7408-7411(1997).
[16]
RN CHARACTERIZATION OF ANGIOSTATIN, AND PARTIAL SEQUENCE.
RX MEDLINE-95042728; PubMed-7525077;
RA O'Reilly M.S., Holmgren L., Shing Y., Chen C., Rosenthal R.A.,
RA Moses M., Lane W.S., Cao Y., Sage E.H., Folkman J.;
RT "Angiostatin: a novel angiogenesis inhibitor that mediates the
suppression of metastases by a Lewis lung carcinoma.";
Cell 79:315-328(1994).
[17]
RN CHARACTERIZATION OF ANGIOSTATIN.
RX MEDLINE-97238710; PubMed-9102221;
RA Sim B.K., O'Reilly M.S., Liang H., Fortier A.H., He W., Madsen J.W.,
RA Lapchovich R., Nacy C.A.;
RT "A recombinant human angiotensin protein inhibits experimental primary
and metastatic cancer.";
Cancer Res. 57:1329-1334(1997).
[18]
RN X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF 374-461.
RX MEDLINE-92031502; PubMed-1657148;
RA Mulichak A.M., Tulinsky A., Ravichandran K.G.;
RT "Crystal and molecular structure of human plasminogen kringles 4
refined at 1.9-A resolution.";
Biochemistry 30:10576-10588(1991).
[19]
RN X-RAY CRYSTALLOGRAPHY (2.25 ANGSTROMS) OF 374-461.
RX MEDLINE-92031503; PubMed-1657149;
RA Wu T.-P., Padmanabhan K., Tulinsky A., Mulichak A.M.;
RT "The refined structure of the epsilon-aminocaproic acid complex of
human plasminogen kringles 4.";
Biochemistry 30:10589-10594(1991).
[20]
RN X-RAY CRYSTALLOGRAPHY (1.67 ANGSTROMS) OF 376-454.
RX Stec B., Yamano A., Whitlow M., Teeter M.M.;
RT "Structure of human plasminogen kringles 4 at 1.68 Angstrom and 277 K.
A possible structural role of disordered residues.";
Acta Crystallogr. D 53:169-178(1997).
[21]
RN X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 102-181.
RX MEDLINE-96180681; PubMed-8611560;
RA Mathews I.I., Vanderhoff-Hanover P., Castellino F.J., Tulinsky A.;
RT "Crystal structures of the recombinant kringles 1 domain of human
plasminogen in complexes with the ligands epsilon-aminocaproic acid
and trans-4-(aminomethyl)cyclohexane-1-carboxylic Acid.";
Biochemistry 35:2567-2576(1996).
[22]
RN X-RAY CRYSTALLOGRAPHY (1.66 ANGSTROMS) OF 480-563.
RX MEDLINE-98198034; PubMed-9521645;
RA Chang Y., Mochalkin I., McCance S.G., Cheng B., Tulinsky A.,
RA Castellino F.J.;
RT "Structure and ligand binding determinants of the recombinant kringles
5 domain of human plasminogen.";
Biochemistry 37:3258-3271(1998).
[23]
RN STRUCTURE BY NMR OF 96-184.
RX MEDLINE-94237157; PubMed-8181475;
RA Rejzante M.R., Llinas M.;
RT "1H-NMR assignments and secondary structure of human plasminogen
kringle 1.";
Eur. J. Biochem. 221:927-937(1994).
[24]
RN STRUCTURE BY NMR OF 96-184.
RX MEDLINE-94237158; PubMed-8181476;
RA Rejzante M.R., Llinas M.;
RT "Solution structure of the epsilon-aminohexanoic acid complex of
human plasminogen kringles 1.";
Eur. J. Biochem. 221:939-949(1994).
[25]
RN STRUCTURE BY NMR OF 183-354.
RX MEDLINE-96194156; PubMed-8652577;
RA Soehndel S., Hu C.-K., Marti D., Affolter M., Schaller J., Llinas M.,

RA Rickli E.E.;
RT "Recombinant gene expression and 1H NMR characteristics of the
kringle (2 + 3) supermodule: spectroscopic/functional individuality
of plasminogen kringles domains.";
RL Biochemistry 35:2357-2364(1996).
[26]
RN STRUCTURE BY NMR OF 374-461.
RX MEDLINE-90219023; PubMed-2157850;
RA Atkinson R.A., Williams R.J.P.;
RT "Solution structure of the kringle 4 domain from human plasminogen by
1H nuclear magnetic resonance spectroscopy and distance geometry.";
RL J. Mol. Biol. 212:541-552(1990).
[27]
RN VARIANTS PHE-374 AND THR-620.
RP
Query Match 40.0%; Score 1042; DB 1; Length 810;
Best Local Similarity 43.8%; Pred. No 2e-72;
Matches 193; Conservative 68; Mismatches 140; Indels 42; Gaps 13;
QY 30 TKK---VNTADQCANRCTRNKGLPFTCKAFVFKARKQCLWFFNFSSGVKKFGEHFD 86
DB 37 TKKQLGAGSIECAKCEDE--EFTCRAFOYHSKEQCVIMAEKRKSSIIIRM--RDVV 92
QY 87 LYENKDYIRNCIIIGKRSYKGTVSITKSGIKCQPSWMIPEHSEFLPSSYRGKDLQENYC 146
DB 93 LFEKKVYLSECKTGNGKNYRGTMSTKNGITCQKWSSTSPHRPFPSPATHPSGLEENYC 152
QY 147 RNPREGEGGPMCFWTSNPEVRYEVCIDIQCSEVECMTCGESYRGIMDHTEGKICQRMWH 206
DB 153 RNPNDPQGPWCYTDPKRYDYCDILECEB--ECMHCSGNYDGKISKTMGLECQAWDS 211
QY 207 QTPHHRKFLPERYPDKGFDNDYCRNPDGQPPWCYTILDPHTRWEYCAIKTCADNTMND 266
DB 212 QSPHAGYIPSKFPKNLKNKYNCRNPDRELPRWCFITDPNKRWELCDIPRCT-----TP 265
QY 267 VPLE--TTECTGOGEGYRGVTNIWNCIPCORWDSQVPEHDMTPENFKCKDLRENYCR 324
DB 266 PPSGPTVQCLKGTGENYRGVAVTVSGHTCQHSQAQTPHTHNTPENFCKNLDENYCR 325
QY 325 NPDGSESPWCFITDPIRNVGYCSQIPNCDMSH-----GDCYRGNGKN 367
DB 326 NPDGKRAPWCHTNSQVRYEYC-KIPSCDSPSVSTEQLAPPELTPVVDQYHGDQS 384
QY 368 YGNLSQTRSLGTSMDKMDLHRHFWF--PDASKLNEYCRNPDADDHGPWCYTG 424
DB 385 YRGTSSTTTTKKCSQSSWMTTP--HRHOKTPENYPNAG-LTMNYCRNPDAD-KGPWCFTT 440
QY 425 NPLIPWDYCPISRCGDTTPTIV 447
DB 441 DPSVRYEYCNLKKCSG-TEASVV 462
RESULT 11
ID PLMN BOVIN STANDARD; PRT; 812 AA.
AC P06868; Q28162;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Plasminogen precursor (EC 3.4.21.7).
GN PLG.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Berglund L., Andersen M.D., Petersen T.E.;
RT "Cloning and characterization of the bovine plasminogen cDNA.";
RL Int. Dairy J. 5:593-603(1995).
RN [2]

DT 01-AUG-1988 (Rel. 08, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Apolipoprotein(a) precursor (BC 3.4.21.-) (Apo(a)) (Lp(a)).
GN LPA.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88039109; PubMed=3670400;
RA McLean J.W., Tomlison J.E., Kuang W.-J., Eaton D.L., Chen E.Y.,
RA Fless G.M., Scanu A.M., Lawn R.M.;
RT "cDNA sequence of human apolipoprotein(a) is homologous to
RT plasminogen.";
RL Nature 330:132-137(1987).
RN [2]
RP SERINE PROTEASE ACTIVITY.
RX MEDLINE=90076123; PubMed=2531657;
RA Salonen E.-M., Jauhainen M., Zardi L., Vaheri A., Ehnholm C.;
RT "Lipoprotein(a) binds to fibronectin and has serine proteinase
RT activity capable of cleaving it.";
RL EMBO J. 8:4035-4040(1989).
RN [3]
RP REVIEW.
RX MEDLINE=90049223; PubMed=2530631;
RA Utermann G.;
RT "The mysteries of lipoprotein(a).";
RL Science 246:904-910(1989).
RN [4]
RP CHARACTERIZATION OF THE N- AND O-LINKED GLYCANS.
RX MEDLINE=21303595; PubMed=11294842;
RA Garner B., Merry A.H., Royle L., Harvey D.J., Rudd P.M., Thillet J.;
RT "Structural elucidation of the N- and O-glycans of human
RT apolipoprotein(a): role of o-glycans in conferring protease
RT resistance.";
RL J. Biol. Chem. 276:22200-22208(2001).
RN [5]
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 4121-4208.
RX MEDLINE=96217891; PubMed=8642595;
RA Mikol V., Lograsso P.V., Boettcher B.R.;
RT "Crystal structures of apolipoprotein(a) kringle IV37 free and
RT complexed with 6-aminohexanoic acid and with p-aminomethylbenzoic
RL acid: existence of novel and expected binding modes.";
RL J. Mol. Biol. 256:751-761(1996).
RN [6]
RP VARIANT ARG-4193.
RX MEDLINE=95002201; PubMed=7918682;
RA Scanu A.M., Pfaffinger D., Lee J.C., Hinman J.;
RT "A single point mutation (Trp72-->Arg) in human apo(a) kringle 4-37
RT associated with a lysine binding defect in Lp(a).";
RL Biochim. Biophys. Acta 1227:41-45(1994).
RN [7]
RP FUNCTION: Apo(a) is the main constituent of lipoprotein(a)
CC (Lp(a)). It has serine proteinase activity and is able of
CC autolysis, inhibits tissue-type plasminogen activator 1.
CC Lp(a) may be a ligand for megalin/Sp 330.
CC [8] SUBUNIT: Disulfide-linked to apo-B100. Binds to fibronectin and
CC decorin.
CC [9] PTM: N- and O-glycosylated. The N-glycans are complex biantennary
CC structures present in either a mono- or disialylated state. The
CC O-glycans are mostly (80%) represented by the monosialylated core
CC type I structure, NeuNAc2GalNAc6, with smaller
CC amounts of disialylated and non-sialylated O-glycans also
CC detected.
CC [10] DISEASE: Elevated plasma concentrations of apo(a) and its
CC naturally occurring proteolytic fragments are correlated with
CC atherosclerosis. Homology with plasminogen kringle IV and V is
CC thought to underlie the atherogenicity of the protein, because the
CC fragments are competing with plasminogen for fibrin(ogen) binding.
CC [11] MISCELLANEOUS: Apo(a) is known to be proteolytically cleaved,
CC leading to the formation of the so called mini-Lp(a). Apo(a)
CC fragments accumulate in atherosclerotic lesions, where they may
CC promote thrombogenesis. O-glycosylation may limit the extent of
CC

CC proteolytic fragmentation.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. PLASMINOGEN SUBFAMILY.
CC -!- SIMILARITY: CONTAINS 38 KRINGLE DOMAINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: X06290; CAA29618.1; .
CC PIR: S00657; S00657.
CC HSP: P00747; IPMK.
CC MEROPS: S01.226; .
CC Genew: HGNC:6667; LPA.
CC MIM: 152200; .
CC InterPro: IPR001314; Chymotrypsin.
CC InterPro: IPR000001; Kringle.
CC InterPro: IPR001254; Ser.protease_Try.
CC Pfam: PF00051; Kringle; 38.
CC Pfam: PF00089; trypsin; 1.
CC PRINTS: PR00722; CHYMOTRYPSIN.
CC PRINTS: PR00018; KRINGLE.
CC ProDom: PD000395; Kringle; 38.
CC SMART: SM00130; KR; 38.
CC SMART: SM00020; Tryp_Spc; 1.
CC PROSITE: PS00021; KRINGLE_1; 38.
CC PROSITE: PS50070; KRINGLE_2; 38.
CC PROSITE: PS50240; TRYPSIN_DOM; 1.
CC PROSITE: PS00134; TRYPSIN_HIS; 1.
CC PROSITE: PS00135; TRYPSIN_SER; 1.
CC Hydrolase: Serine protease; Lipid transport; Plasma; Glycoprotein;
CC Kringle; Repeat; Atherosclerosis; Signal; Polymorphism.
CC
CC SIGNAL 1 19
CC CHAIN 20 4548 APOLIPOPROTEIN(A).
CC DOMAIN 20 130 KRINGLE TYPE IV, 1.
CC DOMAIN 131 244 KRINGLE TYPE IV, 2.
CC DOMAIN 245 358 KRINGLE TYPE IV, 3.
CC DOMAIN 359 472 KRINGLE TYPE IV, 4.
CC DOMAIN 473 586 KRINGLE TYPE IV, 5.
CC DOMAIN 587 700 KRINGLE TYPE IV, 6.
CC DOMAIN 701 814 KRINGLE TYPE IV, 7.
CC DOMAIN 815 928 KRINGLE TYPE IV, 8.
CC DOMAIN 929 1042 KRINGLE TYPE IV, 9.
CC DOMAIN 1043 1156 KRINGLE TYPE IV, 10.
CC DOMAIN 1157 1270 KRINGLE TYPE IV, 11.
CC DOMAIN 1271 1384 KRINGLE TYPE IV, 12.
CC DOMAIN 1385 1498 KRINGLE TYPE IV, 13.
CC DOMAIN 1499 1612 KRINGLE TYPE IV, 14.
CC DOMAIN 1613 1726 KRINGLE TYPE IV, 15.
CC DOMAIN 1727 1840 KRINGLE TYPE IV, 16.
CC DOMAIN 1841 1954 KRINGLE TYPE IV, 17.
CC DOMAIN 1955 2068 KRINGLE TYPE IV, 18.
CC DOMAIN 2069 2182 KRINGLE TYPE IV, 19.
CC DOMAIN 2183 2296 KRINGLE TYPE IV, 20.
CC DOMAIN 2297 2410 KRINGLE TYPE IV, 21.
CC DOMAIN 2411 2524 KRINGLE TYPE IV, 22.
CC DOMAIN 2525 2638 KRINGLE TYPE IV, 23.
CC DOMAIN 2639 2752 KRINGLE TYPE IV, 24.
CC DOMAIN 2753 2866 KRINGLE TYPE IV, 25.
CC DOMAIN 2867 2980 KRINGLE TYPE IV, 26.
CC DOMAIN 2981 3094 KRINGLE TYPE IV, 27.
CC DOMAIN 3095 3208 KRINGLE TYPE IV, 28.
CC DOMAIN 3209 3322 KRINGLE TYPE IV, 29.
CC DOMAIN 3323 3436 KRINGLE TYPE IV, 30.
CC DOMAIN 3437 3550 KRINGLE TYPE IV, 31.
CC DOMAIN 3551 3664 KRINGLE TYPE IV, 32.
CC DOMAIN 3665 3770 KRINGLE TYPE IV, 33.
CC DOMAIN 3771 3884 KRINGLE TYPE IV, 34.
CC DOMAIN 3885 3998 KRINGLE TYPE IV, 35.
CC DOMAIN 3999 4112 KRINGLE TYPE IV, 36.
CC

```

FT DOMAIN 4113 4226 KRINGLE TYPE IV, 37.
FT DOMAIN 4227 4327 KRINGLE TYPE V.
FT DOMAIN 4328 4548 SERINE PROTEASE.
FT ACT_SITE 4369 4369 CHARGE RELAY SYSTEM.
FT ACT_SITE 4412 4412 CHARGE RELAY SYSTEM.
FT ACT_SITE 4498 4498 CHARGE RELAY SYSTEM.
FT VARIANT 4193 4193 W -> R (LOSS OF LYSINE-SEPHAROSE BINDING).
FT SEQUENCE 4548 AA; 501313 MW; 96921BE96A465C5F CRC64;

Query Match 33.8%; Score 879; DB 1; Length 4548;
Best Local Similarity 39.2%; Pred. No. 4e-59;
Matches 166; Conservative 43; Mismatches 138; Indels 76; Gaps 7;

Qy 89 ENKDYIRNCLIGKRSYKGTVSITKGIKOPWSSMIPHEHSFLPSSYRGKDLQENYCRN 148
Db 3888 ENSTGVQDVRGDSYRGTLSTITGRTCSWSSMTPHRRIRLYPNAGLTRYCRN 3947

Qy 149 PRGEGGPGWCFSTSNPEVRYEYCDIPQCSEV-----
Db 3948 PDAEI-RPWCYTMDPSYRWEYCNLTRCPVTESVLTTPVAVPSTEAPSEQAPEKSPV 4006

Qy 180 ---CWTNGESYRGLMDHTESGKTCQRWDHOTPHRHFLPERYPDKGFDNDCYCRNPQGP 236
Db 4007 VQDCYHGDGRSYRGSISSTVTGRTCSWSSMIPHWHQRTPENYPNAGLTENYCRNPDSGK 4066

Qy 237 RPWCYTLDPHTRWEYCAIKTCADNTM---NDTDVPLETTE-----CIQ 276
Db 4067 QPWCYTLDPCVRWEYCNLTQSETSEVLETPVTPVPSMEAHSEAAPEQTGTPVVRQCYH 4126

Qy 277 GQSGYRGTVNTIWNIGPCQRWDSDQYRPHEDMTPEFKCKDLRENYCRNPDSGSPWCFT 336
Db 4127 GNGOSYRGTFSTVTGRTCSWSSMTPHRHQRTPENYPNDGLTWNCRNPADTGPWCFT 4186

Qy 337 TDPNIRVGYCS-----QIPNCDMSHGQDCYRGNGKYNMGNLSQTRSG 378
Db 4187 MDPSIRWEYCNLTRCSDTEGTVAAPTVIQVPSLGPPSEQDCMFGNGKGYRGKATTVTG 4246

Qy 379 LTCSMDKNMDELHRLFWPEADSK---LNEYCRNPDDDAHGPWCYTGPNLIPWDYCP 435
Db 4247 TPCQEW-AAQEPHRHSTFTPGTNKAGLEKNYCRNPDDGNGPWCYTMNPKLFEDYCDI 4304

Qy 436 SRC 438
Db 4305 PLC 4307

RESULT 13
APOA_MACMU STANDARD; PRT; 1420 AA.
AC P1417;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Apo(a)lipoprotein(a) (EC 3.4.21.-) (Apo(a)) (Lp(a)) (Fragment).
GN Lp(a).
OS Macaca mulatta (Rhesus macaque).
OC Eukaryota; Metazoa; Chordata; Graniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecoidea; Macaca.
OX NCBI_TaxID=9544;
RN [1]
RP SEQUENCE FROM N.A.
RA Tomlinson J.E., McLean J.W., Lawn R.M.;
RT "Rhesus monkey apolipoprotein(a). Sequence, evolution, and sites of
RL J. Biol. Chem. 264:5957-5965(1989).
CC -1- FUNCTION: Apo(a) is the main constituent of lipoprotein(a)
CC (Lp(a)). It has serine proteinase activity and is able of
CC autoproteolysis. Inhibits tissue-type plasminogen activator 1.
CC Lp(a) may be a ligand for megalin/Gp 330.

```

```

-1- SUBUNIT: Disulfide-linked to apo-B100. Binds to fibronectin and
decorin (By similarity).
-1- PTM: N- and O-glycosylated (By similarity).
-1- DISEASE: Elevated plasma concentrations of apo(a) and its
naturally occurring proteolytic fragments are correlated with
atherosclerosis. Homology with plasminogen kringle IV and V is
thought to underlie the atherogenicity of the protein, because the
fragments are competing with plasminogen for fibrin(ogen) binding.
-1- MISCELLANEOUS: Apo(a) is known to be proteolytically cleaved,
leading to the formation of the so called mini-Lp(a). Apo(a)
fragments accumulate in atherosclerotic lesions, where they may
promote thrombogenesis. O-glycosylation may limit the extent of
proteolytic fragmentation (By similarity).
-1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. PLASMINOGEN SUBFAMILY.
-1- SIMILARITY: CONTAINS AT LEAST 10 KRINGLE DOMAINS.
-----
THIS SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; J04635; AAA36833.1;
PIR; A30848; A30848.
PIR; A32869; A32869.
HSP; P00747; 2PK4.
MEROPS; S01.226; -.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00051; kringle; 11.
DR Pfam; PF00089; trypsin; 1.
DR ProDom; PD000395; Kringle; 10.
DR SMART; SM00130; KR; 10.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00021; KRINGLE_1; 9.
DR PROSITE; PS00070; KRINGLE_2; 10.
DR PROSITE; PS0240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; FALSE_NEG.
DR PROSITE; PS00135; TRYPSIN_SER; FALSE_NEG.
KW Hydrolase; Serine protease; Lipid transport; Plasma; Glycoprotein;
FT Kringle; Repeat; Atherosclerosis.
FT NON_TER 1
FT DOMAIN 49 127 KRINGLE 1.
FT DOMAIN 163 241 KRINGLE 2.
FT DOMAIN 277 355 KRINGLE 3.
FT DOMAIN 391 469 KRINGLE 4.
FT DOMAIN 505 583 KRINGLE 5.
FT DOMAIN 619 697 KRINGLE 6.
FT DOMAIN 725 803 KRINGLE 7.
FT DOMAIN 839 917 KRINGLE 8.
FT DOMAIN 953 1031 KRINGLE 9.
FT DOMAIN 1067 1145 KRINGLE 10.
FT DOMAIN 1191 1420 SERINE PROTEASE.
SQ SEQUENCE 1420 AA; 158367 MW; BE102949E03C5B0E CRC64;

Query Match 32.4%; Score 843.5; DB 1; Length 1420;
Best Local Similarity 37.2%; Pred. No. 5.9e-57;
Matches 161; Conservative 43; Mismatches 141; Indels 88; Gaps 9;

Qy 89 ENKDYIRNCLIGKRSYKGTVSITKGIKOPWSSMIPHEHSFLPSSYRGKDLQENYCRN 148
Db 270 EQRSGVQECYHNGQSYRGTYFTVTGTCQAWSSMKPHSHSRTPENYPNGGLIRNYCRN 329

Qy 149 PRGEGGPGWCFSTSNPEVRYEYCDIPQCSEV-----
Db 330 P-DPVAAPCYTMDPNVRWEYCNLTQCSDAEGTAVAPNVTPVPSLEAPSEQAPEQRLG 388

Qy 179 --ECWTCNGESYRGLMDHTESGKTCQRWDHOTPHRHFLPERYPDKGFDNDCYCRNPQGP 236
Db 389 VOECYHNSGQSYRGTYFTVTGRTCSWSSMTPHSHSRTPENYPNAGLVKNYCRNPDPVA 448

```

```
QY 237 RPWCYTLDPHTWEYCAIKTCADNTMDTDV-----PLETT-----EC 274
||||| || ||||| : ||| ||| ||| |||
Db 449 APWCYTTDPSVRWEYCNLTRCD--AETAVMPNIIIVPSLEAFLEQETTEETPGVQEC 506
||||| || ||||| : ||| ||| ||| |||
QY 275 IQCQGEYGRYNTIWIWGPCQWDSQYPHEHDMTPENFKCKDLRENYCRNPDSGESPPWC 334
||| : ||||| : ||| ||| ||| : ||| ||| ||| |||
Db 507 YYHGYSGYRGTYSTVTGTCTQAWSSMTPHQHSRTPKYNPNAGLTRNYCRNPDAEIRPWC 566
||||| || ||||| : ||| ||| ||| |||
QY 335 FTTDPNIRVYCS-----QIPNCDSHGDCYRGNGK 366
||| : ||| : |||
Db 567 YTMDSVRWEYCNLTQCLVTSSVLETLTVDPDQTSASSEAPTEQSPVDCYHGGQ 626
||||| || ||||| : ||| ||| ||| |||
QY 367 NYMGNLSQTRSLGTCMMDKNMEDLH-RHIFWEPDASKLNENYCRNPDDAHGPMCTYGN 425
||| : ||| : ||| ||| ||| : ||||| :
Db 627 SYRGSFTVTGTCSNWSMTPHHQRTTEYYPDGG-LTRNYCRNPDAEIR-PWCYTM 684
||||| || ||||| : ||| ||| ||| |||
QY 426 PLIPWDYCPISRC 438
||| : ||| : |||
Db 685 PSVRWEYCNLTQC 697
||||| || ||||| : ||| ||| ||| |||

RESULT 14
PLMN_PETMA STANDARD; PRT; 325 AA.
AC P33574;
DF 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Plasminogen (EC 3.4.21.7) (Fragments).
OS Petromyzon marinus (Sea lamprey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Hypercoartia;
OC Petromyzontiformes; Petromyzontidae; Petromyzon.
OX NCBI_TaxID=7757;
RN [1]
RP SEQUENCE.
RA Afolter M., Schaller J., Rickli E.E.;
RT "Isolation, characterization and partial amino acid sequence of
RT lamprey plasminogen.";
RL Protein Seq. Data Anal. 5:207-211(1993).
CC -1- FUNCTION: PLASMIN DISSOLVES THE FIBRIN OF BLOOD CLOTS AND ACTS AS
CC A PROTEOLYTIC FACTOR IN A VARIETY OF OTHER PROCESSES INCLUDING
CC EMBRYONIC DEVELOPMENT, TISSUE REMODELING, TUMOR INVASION,
CC AND INFLAMMATION.
CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Lys-|-Xaa > Arg-|-Xaa;
CC higher selectivity than trypsin. Converts fibrin into soluble
CC products.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. PLASMINOGEN SUBFAMILY.
CC -1- SIMILARITY: CONTAINS AT LEAST 2 KRINGLE DOMAINS.
CC PIR; S33879; S33879.
DR HSP; P00747; SHPG.
DR MEROPS; S01.233; -.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_protease_Try.
DR PRINTS; PR00018; KRINGLE.
DR PRODOM; PD000395; Kringle; 2.
DR SMART; SM00130; KR; 2.
DR PROSITE; PS00021; KRINGLE_1; 2.
DR PROSITE; PS00070; KRINGLE_2; 2.
DR PROSITE; PS00240; TRYPsin DOM; PARTIAL.
DR PROSITE; PS00134; TRYPsin HIS; PARTIAL.
DR PROSITE; PS00135; TRYPsin SER; PARTIAL.
KW Hydrolase; Serine protease; Plasma; Glycoprotein; Fibrinolysis;
KW Tissue remodeling; Blood coagulation; Kringle; Zymogen.
FT NON_CONS 15 16
FT NON_CONS 34 35
FT NON_CONS 44 45
FT NON_CONS 76 77
FT NON_CONS 111 112
FT NON_CONS 138 139
FT NON_CONS 158 159
FT NON_CONS 178 179
FT NON_CONS 216 217
FT NON_CONS 236 237
```

```
FT NON_CONS 267 268
FT NON_CONS 282 283
FT NON_CONS 295 296
FT NON_CONS 307 308
FT NON_CONS 315 316
FT NON_TER 325 325
SQ SEQUENCE 325 AA; 35194 MW; 1B5F0B539AC6ED3C CRC64;

Query Match 15.7%; Score 408.5; DB 1; Length 325;
Best Local Similarity 28.7%; Pred. No. 2.4e-24;
Matches 98; Conservative 33; Mismatches 68; Indels 143; Gaps 12;

QY 106 KG-TVSTTKSIKQDPWSSMIPHEHDMTPENFKCKDLRENYCRNPDSGESPPWCFTSNPE 164
||| : ||| : ||| ||| : ||| ||| ||| |||
Db 4 KGYSVTVXLYIFDCQKWSNYPKPNFSDAT-----DPK-----GPMCYTTD-- 45
||||| || ||||| : ||| ||| ||| |||
QY 165 VRYEVCIDPQCSVECECTCGESYRGLMDHTESGKICQWRDHPHHRKFLPERYPKGF 224
||| : ||| : ||| ||| ||| : ||| ||| ||| |||
Db 46 -----YXGAASVTRSG---LRGDEQTPHRTFSPQSFAGL-- 77
||||| || ||||| : ||| ||| ||| |||
QY 225 DDNYCRNPDGQPRPWCYTLDPHTRWEYCAIKTCADNTMDTDVPLETTECIQGGEGYRG 284
||||| || ||||| : ||| ||| ||| |||
Db 78 -----TTACVKGTGEYRG 91
||||| || ||||| : ||| ||| ||| |||
QY 285 TVNTIWIWGPCQWDSQYPHEHDMTPENFKCKDLRENYCRNPDSGESPPWCFTTNIRVG 344
||| : ||| ||| ||| : ||| ||| ||| ||| : ||| |||
Db 92 TAALTVSQKACQAWASQTP-----GDVYSCQGLVSNYCRNPDKGKLPWCYTTE----- 139
||||| || ||||| : ||| ||| ||| |||
QY 345 YCSQIPNCD-----MSHGQDCYRGNGKNGMNLQTSRSLGTCSMWDMKNMEDLHR 393
||| : ||| : ||| ||| ||| : ||| ||| ||| |||
Db 140 YCN-VPSTGTGPTSEYHEILTPTAQDXYTGIVEDYRGKMS----- 178
||||| || ||||| : ||| ||| ||| |||
QY 394 HIFWEPDASKLNENYCRNPDDAHGPMCTYGNPLIPWDYCP 435
||||| || ||||| : ||| ||| ||| |||
Db 179 -----PDAG-LEENFCRNPDDQPGQPMCYTXNPEAXPRYCDV 214
||||| || ||||| : ||| ||| ||| |||

RESULT 15
THRB_BOVIN STANDARD; PRT; 625 AA.
ID THRB_BOVIN AC P00735;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Prothrombin precursor (EC 3.4.21.5).
GN F2.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88245190; PubMed=3379642;
RA Irwin D.M., Robertson K.A., Macgillivray R.T.A.;
RT "Structure and evolution of the bovine prothrombin gene.";
RL J. Mol. Biol. 200:31-45(1988).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=84203525; PubMed=6326805;
RA McGillivray R.T.A., Davie E.W.;
RT "Characterization of bovine prothrombin mRNA and its translation
RT product.";
RL Biochemistry 23:1626-1634(1984).
RN [3]
RP SEQUENCE OF 44-625, DISULFIDE BONDS, AND CARBOHYDRATE-LINKAGE SITES.
RA Magnusson S., Sottrup-Jensen L., Petersen T.E., Claess H.;
RL (In) Hemker H.C., Veltkamp J.J. (eds.);
RL Boerhaave symposium on prothrombin and related coagulation factors,
RL pp.25-46, Leiden University Press, Leiden (1975).
RN [4]
RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS) OF ACTIVATION PEPTIDE 1.
RX MEDLINE=86296631; PubMed=3741841;
```

RA Park C.H., Tulinsky A.;
 RT "Three-dimensional structure of the kringle sequence: structure of
 RT prothrombin fragment 1.";
 RL Biochemistry 25:3977-3982(1986).
 [5]
 RN X-RAY CRYSTALLOGRAPHY (2.25 ANGSTROMS) OF ACTIVATION PEPTIDE 1.
 RP MEDLINE-91311686; PubMed-1856869;
 RX Seshadri T.-P., Tulinsky A., Skrzypczak-Jankun E., Park C.H.;
 RA "Structure of bovine prothrombin fragment 1 refined at 2.25-A
 RT resolution.";
 RL J. Mol. Biol. 220:481-494(1991).
 [6]
 RN X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF ACTIVATION PEPTIDE 1.
 RP MEDLINE-92190185; PubMed-1547238;
 RX Soriano-Garcia M., Padmanabhan K., de Vos A.M., Tulinsky A.;
 RA "The Ca2+ ion and membrane binding structure of the Gla domain of Ca-
 RT prothrombin fragment 1.";
 RL Biochemistry 31:2554-2566(1992).
 [7]
 RN X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS).
 RP MEDLINE-92218459; PubMed-1560020;
 RX Martin P.D., Robertson W., Turk D., Huber R., Bode W., Edwards B.F.P.;
 RA "The structure of residues 7-16 of the A alpha-chain of human
 RT fibrinogen bound to bovine thrombin at 2.3-A resolution.";
 RL J. Biol. Chem. 267:7911-7920(1992).
 [8]
 RN X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS).
 RP MEDLINE-92389319; PubMed-1518046;
 RX Brandstetter H., Turk D., Hoeffken H.W., Grosse D., Stuerzebecher J.,
 RA Martin P.D., Edwards B.F.P., Bode W.;
 RT "Refined 2.3 A X-ray crystal structure of bovine thrombin complexes
 RT formed with the benzamidine and arginine-based thrombin inhibitors
 RT NAPAP, 4-TAPAP and MQPA. A starting point for improving
 RT antithrombotics.";
 RL J. Mol. Biol. 226:1085-1089(1992).
 [9]
 RN X-RAY CRYSTALLOGRAPHY (3.1 ANGSTROMS) OF COMPLEX WITH ORNITHODORIN.
 RP MEDLINE-97102783; PubMed-8947023;
 RX van de Loch A., Stubbs M.T., Bode W., Friedrich T., Bollschweiler C.,
 RA Hoeffken W., Huber R.;
 RT "The ornithodorin-thrombin crystal structure, a key to the TAP
 RT enigma?";
 RL EMBO J. 15:6011-6017(1996).
 [10]
 RN X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF COMPLEX WITH TRIABIN.
 RP MEDLINE-98004486; PubMed-9342325;
 RX Fuentes-Prior P., Noeske-Jungblut C., Donner P., Schleuning W.D.,
 RA Huber R., Bode W.;
 RT "Structure of the thrombin complex with triabin, a lipocalin-like
 RT exosite-binding inhibitor derived from a triatomine bug.";
 RL Proc. Natl. Acad. Sci. U.S.A. 94:11845-11850(1997).
 [11]
 RN GENE STRUCTURE.
 RP MEDLINE-86077733; PubMed-3000440;
 RX Irwin D.M., Ahern K.G., Pearson G.D., McGillivray R.T.A.;
 RA "Characterization of the bovine prothrombin gene.";
 RL Biochemistry 24:6854-6861(1985).
 CC -1- FUNCTION: THROMBIN, WHICH CLEAVES BONDS AFTER ARG & LYS, CONVERTS
 CC FIBRINOGEN TO FIBRIN AND ACTIVATES FACTORS V, VII, VIII, XIII,
 CC AND, IN COMPLEX WITH THROMBOMODULIN, PROTEIN C.
 CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Arg-I-Gly; activates
 CC fibrinogen to fibrin and releases fibrinopeptide A and B.
 CC -1- SUBCELLULAR LOCATION: Extracellular.
 CC -1- TISSUE SPECIFICITY: SYNTHESIZED IN THE LIVER; FOUND IN PLASMA.
 CC RESULT FROM THE CARBOXYLATION OF GLUTAMYL RESIDUES, WHICH BIND CALCIUM IONS,
 CC ENZYME, THE VITAMIN K-DEPENDENT CARBOXYLASE. THE MODIFIED RESIDUES
 CC ARE NECESSARY FOR THE CA-DEPENDENT INTERACTION WITH A NEGATIVELY
 CC CHARGED PHOSPHOLIPID SURFACE, WHICH IS ESSENTIAL FOR THE CONVERSION
 CC OF PROTHROMBIN TO THROMBIN.
 CC -1- MISCELLANEOUS: PROTHROMBIN IS ACTIVATED ON THE SURFACE OF A
 CC PHOSPHOLIPID MEMBRANE THAT BINDS THE AMINO END OF PROTHROMBIN &
 CC FACTORS VA & XA IN CA-DEPENDENT INTERACTIONS; FACTOR XA REMOVES

CC THE ACTIVATION PEPTIDE & CLEAVES THE REMAINING PART INTO LIGHT &
 CC HEAVY CHAINS. THE ACTIVATION PROCESS STARTS SLOWLY BECAUSE FACTOR
 CC V ITSELF HAS TO BE ACTIVATED BY THE INITIAL, SMALL AMOUNTS OF
 CC THROMBIN.
 CC -1- MISCELLANEOUS: THROMBIN CAN ITSELF CLEAVE THE AMINO TERMINAL
 CC FRAGMENT (FRAGMENT 1) OF THE PROTHROMBIN, PRIOR TO ITS ACTIVATION
 CC BY FACTOR XA.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 CC -1- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.
 CC -1- DATABASE: NAME-Prozyme technical fact sheet;
 CC WWW="http://www.prozyme.com/technical/thrombindata.html".
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see http://www.isb-sib.ch/announcement/
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; V00135; CAA23451.1; -
 CC EMBL; J00041; AAA30781.1; -
 CC PIR; A00915; TBBO
 CC PIR; S02537; S02537.
 CC PDB; 1B8R; 31-JAN-94.
 CC PDB; 1ETR; 31-JAN-94.
 CC PDB; 1ETS; 31-JAN-94.
 CC PDB; 1ETT; 31-JAN-94.
 CC PDB; 1H8T; 31-JAN-94.
 CC PDB; 2PFI; 31-JAN-94.
 CC PDB; 2PF2; 31-JAN-94.
 CC PDB; 2SPT; 31-MAY-94.
 CC PDB; 1MKX; 07-JUL-97.
 CC PDB; 1MKV; 07-JUL-97.
 CC PDB; 1TBQ; 14-OCT-96.
 CC PDB; 1TBR; 14-OCT-96.
 CC PDB; 1TOC; 23-JUL-97.
 CC PDB; 1VIT; 21-APR-97.
 CC PDB; 1YCP; 06-MAY-98.
 CC PDB; 1A0H; 17-JUN-98.
 CC PDB; 1AVG; 16-FEB-99.
 CC MEROPS; S01.217; -
 CC InterPro; IPR001314; Chymotrypsin.
 CC InterPro; IPR002383; GLA_blood.
 CC InterPro; IPR000001; Kringle.
 CC InterPro; IPR003966; Prothrombin.
 CC InterPro; IPR001254; Ser_protease_Try.
 CC InterPro; IPR000294; VitK_dep_GLA.
 CC Pfam; PF00051; Kringle; 2.
 CC Pfam; PF00089; trypsin; 1.
 CC Pfam; PF00594; gla; 1.
 CC PRINTS; PR00722; CHYMOTRYPSIN.
 CC PRINTS; PR00001; GLABLOOD.
 CC PRINTS; PR00018; KRINGLE.
 CC PRINTS; PR01505; PROTHROMBIN.
 CC ProDom; PD000395; Kringle; 2.
 CC SMART; SM00069; GLA; 1.
 CC SMART; SM00130; KR; 2.
 CC SMART; SM00020; Tryp_SPC; 1.
 CC PROSITE; PS00011; GLU_CARBOXYLATION; 1.
 CC PROSITE; PS00021; KRINGLE_1; 2.
 CC PROSITE; PS00070; KRINGLE_2; 2.
 CC PROSITE; PS02040; TRYPSIN_DOM; 1.
 CC PROSITE; PS00134; TRYPSIN_HIS; 1.
 CC PROSITE; PS00135; TRYPSIN_SER; 1.
 CC Blood coagulation; Plasma; Calcium-binding; Glycoprotein; Repeat;
 CC Vitamin K; Zymogen; Gamma-carboxyglutamic acid; Acute phase; Liver;
 CC Hydrolyase; Serine protease; Kringle; Signal; 3D-structure.
 CC SIGNAL 1 24 POTENTIAL..
 CC PROPEP 1 25 43
 CC CHAIN 44 625.
 CC PEPTIDE 44 199
 CC PEPTIDE 200 317
 CC ACTIVATION PEPTIDE (FRAGMENT 1).
 CC ACTIVATION PEPTIDE (FRAGMENT 2).

```
FT CHAIN 318 366 THROMBIN LIGHT CHAIN (A).
FT CHAIN 367 625 THROMBIN HEAVY CHAIN (B).
FT DOMAIN 109 187 KRINGLE 1.
FT DOMAIN 214 292 KRINGLE 2.
FT DOMAIN 367 625 SERINE PROTEASE.
FT SITE 199 200 CLEAVAGE (BY THROMBIN).
FT SITE 317 318 CLEAVAGE (BY FACTOR XA).
FT SITE 366 367 CLEAVAGE (BY FACTOR XA).
FT SITE 409 409 CHARGE RELAY SYSTEM.
FT ACT_SITE 465 465 CHARGE RELAY SYSTEM.
FT ACT_SITE 571 571 CHARGE RELAY SYSTEM.
FT MOD_RES 50 50 GAMMA-CARBOXYGLUTAMIC ACID.
FT MOD_RES 51 51 GAMMA-CARBOXYGLUTAMIC ACID.
FT MOD_RES 58 58 GAMMA-CARBOXYGLUTAMIC ACID.
FT MOD_RES 60 60 GAMMA-CARBOXYGLUTAMIC ACID.
FT MOD_RES 63 63 GAMMA-CARBOXYGLUTAMIC ACID.
FT MOD_RES 64 64 GAMMA-CARBOXYGLUTAMIC ACID.
FT MOD_RES 69 69 GAMMA-CARBOXYGLUTAMIC ACID.
FT MOD_RES 70 70 GAMMA-CARBOXYGLUTAMIC ACID.
FT MOD_RES 73 73 GAMMA-CARBOXYGLUTAMIC ACID.

Query Match 13.7%; Score 357.5; DB 1; Length 625;
Best Local Similarity 29.9%; Pred. No. 3.9e-20;
Matches 92; Conservative 39; Mismatches 114; Indels 63; Gaps 11;

Qy 172 IPOCSEVETCNGESYRGLMDHTESGKICORWDHQTPHRHKFLPERYPDKGDDNYCRN 231
Db 9 LPGCLALALSLVHSQHVFLAHQAASSLLQR-----ARRANGFLEEVRK- 54

Qy 232 PDGQPRPWCYTLDPHTRWE-----YCAIKTCADNTMNDTDVPLETTE--CIQ 276
Db 55 --GNLERECLE-EPCSRREEAFEALESLSATDAFWAKYFACESARNPREKNECLEGNCAE 111

Qy 277 GQEGYRGTVNTIWIQPCQWDSQYPHEHDMTPENFKCDLRENYCRNPDGS-ESPWCF 335
Db 112 GVGMYRGVSVTRSGIECOLWRSRPHKPEINSTTHPCADLRENFRCRNPDSGITGPWCY 171

Qy 336 TTDPNIRVGYCSQIPNCDSHSGQD-----CYRGNKGNYMGNL 372
Db 172 TTSPTLRRECS-VFVC---GQDRVTVEIPRSGGSTTSQSPLLETCPDRGREYRGRL 226

Qy 373 SOTRSLGTCSMW-DKNMEDLHRHIFWEPDASKLNEYCRNPDDDDAHGPWCYTGNPLIPWD 431
Db 227 AVTTSGRCLAWSSEQAKLSKQDFNP-AVPLAENFCRNPDPGDEGAWCYVADQPGDFE 285

Qy 432 YCPISRCE 439
Db 286 YCDLNYCE 293
```

Search completed: April 25, 2003, 15:51:03
Job time : 32 secs

GenCore version 5.1.4_p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 25, 2003, 15:48:06 ; Search time 35 seconds
(without alignments)
2631.516 Million cell updates/sec

Title: us-09-674-377b-1

Perfect score: 2604

Sequence: 1 ERRRRNTIHEFKKSAKTLI.....IPWDYCPISRCGGTPTTIV 447

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_21:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mmc:*
8: sp_organelle:*
9: sp_phase:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archaea:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2452	94.2	728	6 Q9BH09	Q9bh09 felis silve
2	2053	78.8	726	13 Q90978	Q90978 gallus gall
3	1864.5	71.6	710	13 Q91402	Q91402 xenopus. he
4	1471	56.5	290	4 Q02935	Q02935 homo sapien
5	1463	56.2	296	4 Q14519	Q14519 homo sapien
6	1433.5	55.0	285	4 Q8TCE2	Q8tce2 homo sapien
7	1209.5	46.4	704	13 Q90865	Q90865 gallus gall
8	1208	46.4	716	13 Q91691	Q91691 xenopus lae
9	1176	45.2	717	13 P70006	P70006 xenopus lae
10	1172	45.0	709	13 Q902N6	Q902n6 brachydanio
11	1163.5	44.7	567	4 Q13208	Q13208 homo sapien
12	1144	43.9	716	11 P70521	P70521 rattus norv
13	1125	43.2	716	11 Q91XG8	Q91xc8 mus musculu
14	1085.5	41.7	812	11 Q9R0W3	Q9r0w3 rattus norv
15	1080.5	41.5	806	6 Q18793	Q18793 macropus eu
16	1064	40.9	812	11 Q91WJ5	Q91wj5 mus musculu

ALIGNMENTS

RESULT 1

Q9BH09 ID Q9BH09 PRELIMINARY; PRT; 728 AA.

AC Q9BH09;

DT 01-JUN-2001 (TREMBlrel. 17, Created)

DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)

DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)

DE Hepatocyte growth factor HGF.

OS Felis silvestris catus (Cat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.

OX NCBI_taxID=9685;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=LIVER;

RA Kobayashi Y., Nakamura N., Ishizaka T., Masuda K., Ohno K.,

RA Tsujimoto H.;

RT "Molecular cloning of feline hepatocyte growth factor (HGF) cDNA."

RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.

CC -!- SIMILARIY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE

CC TRYPSIN FAMILY.

DR EMBL; AB046610; BAB21499.1; -.

DR HSSP; P14210; 1BHT.

DR MEROPS; S01.982; -.

DR InterPro; IPR001314; Chymotrypsin.

DR InterPro; IPR000001; Kringle.

DR InterPro; IPR003014; PAN.

DR InterPro; IPR003609; Pan_app.

DR InterPro; IPR001254; Ser_protease_Try.

DR Pfam; PF00051; kringle; 4.

DR Pfam; PF00024; PAN; 1.

DR Pfam; PF00089; trypsin; 1.

DR PRINTS; PR00722; CHYMOTRYPSIN.

DR PRINTS; PR00018; KRINGLE.

DR ProDom; PD000395; Kringle; 4.

DR SMART; SM00130; KR; 4.

DR SMART; SM00473; PAN_AP; 1.

DR SMART; SM00020; Tryp-Spc; 1.

DR PROSITE; PS00021; KRINGLE_1; UNKNOWN_4.

Q15146 homo sapien
Q42341 gallus gall
Q8WMr1 canis famil
Q9bYm0 homo sapien
Q13494 homo sapien
Q9h1v4 homo sapien
Q28398 erinaceus e
Q55027 mus musculu
Q9n1b8 ovis aries
Q46506 papio hamad
Q9bgn9 bos taurus
Q9ylv6 ephydatia f
Q9pu78 crocodylus
Q91001 gallus gall
Q9ptw7 struthio ca
Q91vp2 mus musculu
Q9pu99 homo sapien
Q45507 papio hamad
Q8sq23 sus scrofa
Q16609 homo sapien
Q9uir5 homo sapien
Q9uir6 homo sapien
Q9uir7 homo sapien
Q9uir8 homo sapien
Q25101 herdmania m
Q28911 macaca fasc
Q07153 torpedo cal
Q9brb6 homo sapien
Q9v6k3 drosophila

-!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE TRYPSIN FAMILY.

CC PROSITE: PS50070; KRINGLE_2; 4.
CC EMBL: X84045; CAA58964.1; -.
DR EMBL: X80131; CAA56430.1; -.
KW HSSP: P14210; 1BHT.
DR InterPro: IPR001314; Chymotrypsin.
DR InterPro: IPR000001; Kringle.
DR InterPro: IPR003014; PAN.
DR InterPro: IPR003609; Pan_app.
DR InterPro: IPR001254; Ser_protease_Try.
DR Pfam: PF00051; kringle; 4.
DR Pfam: PF00024; PAN; 1.
DR Pfam: PF00089; trypsin; 1.
DR PRINTS: PR00722; CHYMOTRYPSIN.
DR PRINTS: PR00018; KRINGLE.
DR ProDom: PD000395; kringle; 4.
DR SMART: SM00130; KR; 4.
DR SMART: SM00473; PAN_AP; 1.
DR SMART: SM00020; TRYPSIN; 1.
DR PROSITE: PS00021; KRINGLE_1; 4.
DR PROSITE: PS50070; KRINGLE_2; 4.
DR PROSITE: PS50240; TRYPSIN_DOM; 1.
KW Alternative splicing: Glycoprotein; Growth factor; Hydrolase; Kringle;
KW Serine protease; Serine protease homolog; Signal.
FT SIGNAL 1 27 POTENTIAL.
FT CHAIN 28 726 HEPATOCYTE GROWTH FACTOR/ SCATTER FACTOR.
SQ SEQUENCE 726 AA; 82913 MW; 5805F048A5766C38 CRC64;

Query Match 78.8%; Score 2053; DB 13; Length 726;
Best Local Similarity 76.6%; Pred. No. 1.7e-182;
Matches 341; Conservative 50; Mismatches 52; Indels 2; Gaps 1;

Qy 3 KRRNTIHEFKKSAKTLIKIDPALKIKTKVNTADOCANCRTRNKGLPFTCKAFVDEKAR 62
Db 30 KRRNPLHDYKKTGELMLIKVNTKLETKLNTTEQCAKCRNKGSLFCKAFADRVDT 89
Qy 63 KQCLWFFPNSMSSGVKKEGHEFDLYENKDYIRNCIIIGKGSYKGTIVTSIKSGIKCPWS 122
Db 90 KRCHWLSFNLTNGVRRKQDFAFLFEKKDYVRNCIIIGKAEYKGTIVTSIKSGIQCAWN 149
Qy 123 SMIPHEHSFLPSSYRGKDLQENYCRNPRGEGGPGWCFSTSNPEYRYEVCIDIPOCSEVEECMT 182
Db 150 SMIPHEHSFLPSSYRGKDLQENYCRNPRGEGGPGWCFSTSNPEYRYEVCIDIPOCSEVEECMT 209
Qy 183 CNGESYRGIMDHTEGKICORWDHQTPIRHKFLPERYPDKGFDNVCNRPDQPRWCYT 242
Db 210 CNGESYRGIMDHTEGKICORWDHQTPIRHKFLPERYPDKGFDNVCNRPDQPRWCYT 269
Qy 243 LDPHTRWEYCAIKTCADNTMDTDVPLETTECIOGQEGYRGTVNTIWNIGIPCORWDSQY 302
Db 270 LDPNTPWECAIKTCADNTMDTDVPLETTECIOGQEGYRGTVNTIWNIGIPCORWDSQY 329
Qy 303 PHEHDMTPENFKCKDLQENYCRNPRGEGGPGWCFSTSNPEYRYEVCIDIPOCSEVEECMT 362
Db 330 PHQHNITPENFKCKDLQENYCRNPRGEGGPGWCFSTSNPEYRYEVCIDIPOCSEVEECMT 389
Qy 363 GNGKNTMGNLSOTRSGLTCSMMKNNEDLHRH--IFWEPDASKLNENYCRNPDHAGPW 420
Db 390 GNGKNTMGNLSOTRSGLTCSMMKNNEDLHRH--IFWEPDASKLNENYCRNPDHAGPW 449
Qy 421 CYTGNPLIPWDYCPISRCGDTTPT 445
Db 450 CYTDDPLIPWDYCPISRCGDTTPT 474

RESULT 3
Q91402 ID Q91402 PRELIMINARY; PRT; 710 AA.
AC Q91402;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Hepatocyte growth factor.

DR PROSITE: PS50070; KRINGLE_2; 4.
DR EMBL: X84045; CAA58964.1; -.
KW HSSP: P14210; 1BHT.
DR InterPro: IPR001314; Chymotrypsin.
DR InterPro: IPR000001; Kringle.
DR InterPro: IPR003014; PAN.
DR InterPro: IPR003609; Pan_app.
DR InterPro: IPR001254; Ser_protease_Try.
DR Pfam: PF00051; kringle; 4.
DR Pfam: PF00024; PAN; 1.
DR Pfam: PF00089; trypsin; 1.
DR PRINTS: PR00722; CHYMOTRYPSIN.
DR PRINTS: PR00018; KRINGLE.
DR ProDom: PD000395; kringle; 4.
DR SMART: SM00130; KR; 4.
DR SMART: SM00473; PAN_AP; 1.
DR SMART: SM00020; TRYPSIN; 1.
DR PROSITE: PS00021; KRINGLE_1; 4.
DR PROSITE: PS50070; KRINGLE_2; 4.
DR PROSITE: PS50240; TRYPSIN_DOM; 1.
KW Alternative splicing: Glycoprotein; Growth factor; Hydrolase; Kringle;
KW Serine protease; Serine protease homolog; Signal.
FT SIGNAL 1 27 POTENTIAL.
FT CHAIN 28 726 HEPATOCYTE GROWTH FACTOR/ SCATTER FACTOR.
SQ SEQUENCE 726 AA; 82913 MW; 5805F048A5766C38 CRC64;

Qy 1 EKRRNTIHEFKKSAKTLIKIDPALKIKTKVNTADOCANCRTRNKGLPFTCKAFVDEK 60
Db 30 KRRNPLHDYKKTGELMLIKVNTKLETKLNTTEQCAKCRNKGSLFCKAFADRVDT 89
Qy 61 ARKQCLWFFPNSMSSGVKKEGHEFDLYENKDYIRNCIIIGKGSYKGTIVTSIKSGIKCP 120
Db 90 ARKQCLWFFPNSMSSGVKKEGHEFDLYENKDYIRNCIIIGKGSYKGTIVTSIKSGIKCP 149
Qy 121 WSMIPHEHSFLPSSYRGKDLQENYCRNPRGEGGPGWCFSTSNPEYRYEVCIDIPOCSEVEEC 180
Db 150 WSMIPHEHSFLPSSYRGKDLQENYCRNPRGEGGPGWCFSTSNPEYRYEVCIDIPOCSEVEEC 209
Qy 181 MTCNGESYRGIMDHTEGKICORWDHQTPIRHKFLPERYPDKGFDNVCNRPDQPRWCYT 242
Db 210 MTCNGESYRGIMDHTEGKICORWDHQTPIRHKFLPERYPDKGFDNVCNRPDQPRWCYT 269
Qy 241 YTLDPHTRWEYCAIKTCADNTMDTDVPLETTECIOGQEGYRGTVNTIWNIGIPCORWDS 300
Db 270 YTLDPHTRWEYCAIKTCADNTMDTDVPLETTECIOGQEGYRGTVNTIWNIGIPCORWDS 329
Qy 301 QYHEDMTPENFCKDLQENYCRNPRGEGGPGWCFSTSNPEYRYEVCIDIPOCSEVEEC 360
Db 330 QYHEDMTPENFCKDLQENYCRNPRGEGGPGWCFSTSNPEYRYEVCIDIPOCSEVEEC 389
Qy 361 YRGKNTMGNLSOTRSGLTCSMMKNNEDLHRHIFWEPDASKLNENYCRNPDHAGPW 420
Db 390 YRGKNTMGNLSOTRSGLTCSMMKNNEDLHRHIFWEPDASKLNENYCRNPDHAGPW 449
Qy 421 CYTGNPLIPWDYCPISRCGDTTPT 445
Db 450 CYTDDPLIPWDYCPISRCGDTTPT 474

RESULT 2
Q90978 ID Q90978 PRELIMINARY; PRT; 726 AA.
AC Q90978; Q90966;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Hepatocyte growth factor / scatter factor precursor.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus...
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=EMBRYO;
RX MEDLINE=96029010; PubMed=7554499;
RA Thery C., Sharpe M.J., Bailey S.J., Stern C.D., Gherardi E.;
RT "Expression of HGF/SF, HGF1/MSP, and c-met suggests new functions
RT during early chick development."
RL Dev. Genet. 17:90-101(1995).
RN [2]
RP SEQUENCE OF 1-409 FROM N.A.
RC STRAIN=WHITE LEHORN, AND RHODE ISLAND RED X LIGHT SUSSEX;
RC TISSUE=EMBRYO;
RX MEDLINE=95237013; PubMed=7720585;
RA Streif A., Stern C.D., Thery C., Ireland G.W., Aparicio S.,
RA Sharpe M.J., Gherardi E.;
RT "A role for HGF/SF in neural induction and its expression in Hensen's
RT node during gastrulation."
RL Development 121:813-824(1995).


```
QY 121 WSSMTPHSHSLPSSYRGKDLQENYCRNPRGEGGPGWCFTSNPEVRYEVCDDIPQCSEVEC 180
DB 152 WSSMTPHSHSLPSSYRGKDLQENYCRNPRGEGGPGWCFTSNPEVRYEVCDDIPQCSEVEC 211
QY 181 MTCNGESYRGLMDHTESGKICQWRDHTQPHRHKFLPERYPDKGFDNNYCRNPDGQPRPWC 240
DB 212 MTCNGESYRGLMDHTESGKICQWRDHTQPHRHKFLPERYPDKGFDNNYCRNPDGQPRPWC 271
QY 241 YTLDPHTRWEYCAIKTC 257
DB 272 YTLDPHTRWEYCAIKTC 288
RESULT 5
Q14519 PRELIMINARY; PRT; 296 AA.
AC Q14519;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Competitive HGF antagonist.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LUNG;
RA Chan A.M.-L., Rubin J.S., Bottaro D.P., Hirschfield D.W., Chedid M.,
RA Aaronson S.A.;
RT "Identification of a Competitive HGF Antagonist Encoded by an
RT Alternative Transcript.";
RL Science 0:0-0(1991):80.1; -.
DR EMBL; M77227; AAA35980.1; -.
DR HSP; P14210; 1BHT.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR003014; PAN.
DR InterPro; IPR003609; Pan_app.
DR Pfam; PF00051; kringle; 2.
DR Pfam; PF00024; PAN; 1.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 2.
DR SMART; SM00130; KR; 2.
DR SMART; SM00473; PAN_AP; 1.
DR PROSITE; PS00021; KRINGLE_1; 2.
DR PROSITE; PS00070; KRINGLE_2; 2.
SQ SEQUENCE 296 AA; 34546 MW; A45E456B87AE03BE CRC64;
Query Match 56.2%; Score 1463; DB 4; Length 296;
Best Local Similarity 98.1%; Pred. No. 4.7e-128;
Matches 256; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
QY 1 ERKRRTIHEFKSAKTTLIKIDPALKIKTKVNTADQCANRCTRNKGLPFTCKAFVFDK 60
DB 32 QRKRRTIHEFKSAKTTLIKIDPALKIKTKVNTADQCANRCTRNKGLPFTCKAFVFDK 91
QY 61 ARKOCLEFPNMSGKVEGFEHFDLYENKDYIRNCIIIGKRSYKGTVSITKSGIKCQP 120
DB 92 ARKOCLEFPNMSGKVEGFEHFDLYENKDYIRNCIIIGKRSYKGTVSITKSGIKCQP 151
QY 121 WSSMIPHEHSLPSSYRGKDLQENYCRNPRGEGGPGWCFTSNPEVRYEVCDDIPQCSEVEC 180
DB 152 WSSMIPHEHSLPSSYRGKDLQENYCRNPRGEGGPGWCFTSNPEVRYEVCDDIPQCSEVEC 206
QY 181 MTCNGESYRGLMDHTESGKICQWRDHTQPHRHKFLPERYPDKGFDNNYCRNPDGQPRPWC 240
DB 207 MTCNGESYRGLMDHTESGKICQWRDHTQPHRHKFLPERYPDKGFDNNYCRNPDGQPRPWC 266
RESULT 7
Q90865 PRELIMINARY; PRT; 704 AA.
AC Q90865;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Hepatocyte growth factor-like/macrophage stimulating protein.
GN HGF/MSP.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RX MEDLINE=96029010; PubMed=7554499;
RA Thery C., Sharpe M.J., Batley S.J., Stern C.D., Gherardi E.;
RT "Expression of HGF/SF, HGF/MSP and c-met suggests new functions
RT during early chick development.";
RL Dev. Genet. 17:90-101(1995).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPsin FAMILY.
CC EMBL; X84043; CAA58862.1; -.
DR YHSP; P00747; ICEA.
```

```
RESULT 6
Q8TCE2 PRELIMINARY; PRT; 285 AA.
AC Q8TCE2;
DT 01-JUN-2002 (TREMBLrel. 21, Created)
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Similar to hepatocyte growth factor (hepapoietin A, scatter
DE factor).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=PLACENTA;
RA Strausberg R.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC022308; AAH22308.1; -.
SQ SEQUENCE 285 AA; 33234 MW; 0A93B073EA86EA61 CRC64;
Query Match 55.0%; Score 1433.5; DB 4; Length 285;
Best Local Similarity 97.7%; Pred. No. 2.5e-125;
Matches 251; Conservative 1; Mismatches 0; Indels 5; Gaps 1;
QY 1 ERKRRTIHEFKSAKTTLIKIDPALKIKTKVNTADQCANRCTRNKGLPFTCKAFVFDK 60
DB 32 QRKRRTIHEFKSAKTTLIKIDPALKIKTKVNTADQCANRCTRNKGLPFTCKAFVFDK 91
QY 61 ARKOCLEFPNMSGKVEGFEHFDLYENKDYIRNCIIIGKRSYKGTVSITKSGIKCQP 120
DB 92 ARKOCLEFPNMSGKVEGFEHFDLYENKDYIRNCIIIGKRSYKGTVSITKSGIKCQP 151
QY 121 WSSMIPHEHSLPSSYRGKDLQENYCRNPRGEGGPGWCFTSNPEVRYEVCDDIPQCSEVEC 180
DB 152 WSSMIPHEHSLPSSYRGKDLQENYCRNPRGEGGPGWCFTSNPEVRYEVCDDIPQCSEVEC 206
QY 181 MTCNGESYRGLMDHTESGKICQWRDHTQPHRHKFLPERYPDKGFDNNYCRNPDGQPRPWC 240
DB 207 MTCNGESYRGLMDHTESGKICQWRDHTQPHRHKFLPERYPDKGFDNNYCRNPDGQPRPWC 266
QY 241 YTLDPHTRWEYCAIKTC 257
DB 267 YTLDPHTRWEYCAIKTC 283
RESULT 7
Q90865 PRELIMINARY; PRT; 704 AA.
AC Q90865;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Hepatocyte growth factor-like/macrophage stimulating protein.
GN HGF/MSP.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RX MEDLINE=96029010; PubMed=7554499;
RA Thery C., Sharpe M.J., Batley S.J., Stern C.D., Gherardi E.;
RT "Expression of HGF/SF, HGF/MSP and c-met suggests new functions
RT during early chick development.";
RL Dev. Genet. 17:90-101(1995).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPsin FAMILY.
CC EMBL; X84043; CAA58862.1; -.
DR YHSP; P00747; ICEA.
```

DR MEROPS; S01.977; --
 DR InterPro: IPR001314; Chymotrypsin.
 DR InterPro: IPR000001; Kringle.
 DR InterPro: IPR003014; PAN.
 DR InterPro: IPR003609; Pan_app.
 DR InterPro: IPR001254; Ser_protease_Try.
 DR Pfam: PF00051; kringle; 4.
 DR Pfam: PF00024; PAN; 1.
 DR Pfam: PF00089; trypsin; 1.
 DR PRINTS; PR00722; CHYMOTRYPSIN.
 DR PRINTS; PR00018; KRINGLE.
 DR ProDom: PD000395; Kringle; 4.
 DR SMART; SM00130; KR; 4.
 DR SMART; SM00473; PAN_AP; 1.
 DR SMART; SM00020; TRYPSIN; 1.
 DR PROSITE; PS00021; KRINGLE_1; 4.
 DR PROSITE; PS00070; KRINGLE_2; 4.
 DR PROSITE; PS0240; TRYPSIN_DOM; 1.
 KW Hydrolyase; Serine protease.
 SQ SEQUENCE 704 AA; 79341 MW; CAB0D8CC41367C37 CRC64;

 Query Match 46.4%; Score 1209.5; DB 13; Length 704;
 Best Local Similarity 46.0%; Pred. No. 5.6e-104;
 Matches 207; Conservative 71; Mismatches 149; Indels 23; Gaps 8;

 QY 5 RNTIHEFKSAKTLIKIDPALKIKTKKYNVADQCANRTRNKGLEPFTCKAFVFDKARKQ 64
 DB 20 RSLPNDQFRLRGTETLRAAPNEPPSAPAHGAQAQCAQRCANRP-----DCRAFHERQSOL 75
 QY 65 CLWFPNSMSSGVKKEFEHFDLYENKDYIRNCLIGKGRSYKGTVSITKSGIKCOPWSSM 124
 DB 76 COLLPSQSRSPARLOKNIHYDLQKDFLRECIVANGSYRGTTRTTERGLRQCWHQAT 135
 QY 125 IPHEHFLPSSYRGKDLQENYCRNPRGEGGPMCFSTSNPEVRYEVCIDIPQCSVEVCMTCN 184
 DB 136 TPDRHFLSLRNG--LEENYCRNPRDRGRGWCYTVDFNVRHQSGIKKCEDAVCMTCN 193
 QY 185 GSYRGMDHTESGKTCQRMWDHQTPIRHKFLPERYPDKGFDNNYCRNPDGQPRPWCYTLD 244
 DB 194 GEDYRGFVDTESGTCQRMWDLQHPHKKHPYHPDKYPEKGLDNNYCRNPDSSSEQPMCYTTD 253
 QY 245 PHTRWYCAIKTCADNTMDTVPLE--TTCTCQGGEGYRGVTWNTIWIWGPCQWDSQYP 303
 DB 254 PALEREFCIRVCKKPR-----PINVTTCYRGKGGYRGVNTVTVSGIPQQRWDAQTL 308
 QY 304 HEHDMTPENFKCKDLRENYCRNPDGSEPMCFSTDPNIRVGYCSQIPNCDMS-HGQDCYR 362
 DB 309 HRHFFVPSKYPCKDLQENYCRNPDGSEAPWCFTTRGMRVAFCFHTRCDDDELDAECYH 368
 QY 363 GNGKNTWGNLSQTRSGLTCSMDKNMEDLHRHI-----FWEPDASKLNYCRNPDODDAH 417
 DB 369 GHGRYGHVSKTRKGIQTCQRMWDATTP-----HVPQISPTTHPEA-HLEKNYCRNPDNDSH 423
 QY 418 GPKCYTGNPLIPWDCYPSRCEGDTTPTIV 447
 DB 424 GPMCYTMDPRTDFDYCAIKPCSGSAPSVL 453

RESULT 8
 Q91691 PRELIMINARY; PRT; 716 AA.
 AC Q91691;
 DT 01-NOV-1996 (TremBLrel. 01, Created)
 DT 01-NOV-1996 (TremBLrel. 01, Last sequence update)
 DT 01-JUN-2002 (TremBLrel. 21, Last annotation update)
 DE Growth factor Livertine.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
 OC Xenopodinae; Xenopus.
 OX NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.

RA Ruiz i Altaba A., Thery C.;
 RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL; U57455; AAB52574.1;
 DR HSP; P00747; ICEA.
 DR InterPro: IPR000001; Kringle.
 DR InterPro: IPR003014; PAN.
 DR InterPro: IPR003609; Pan_app.
 DR InterPro: IPR001254; Ser_protease_Try.
 DR Pfam: PF00051; kringle; 4.
 DR Pfam: PF00024; PAN; 1.
 DR Pfam: PF00089; trypsin; 1.
 DR PRINTS; PR00018; KRINGLE.
 DR ProDom: PD000395; Kringle; 4.
 DR SMART; SM00130; KR; 4.
 DR SMART; SM00473; PAN_AP; 1.
 DR SMART; SM00020; TRYPSIN; 1.
 DR PROSITE; PS00021; KRINGLE_1; 4.
 DR PROSITE; PS00070; KRINGLE_2; 4.
 DR PROSITE; PS0240; TRYPSIN_DOM; 1.
 KW Hydrolyase; Serine protease.
 SQ SEQUENCE 716 AA; 81971 MW; 508376A0E4398798 CRC64;

 Query Match 46.4%; Score 1208; DB 13; Length 716;
 Best Local Similarity 46.7%; Pred. No. 7.9e-104;
 Matches 207; Conservative 75; Mismatches 149; Indels 12; Gaps 6;

 QY 5 RNTIHEFKSAKTLIKIDPALKIKTKKYNVADQCANRTRNKGLEPFTCKAFVFDKARKQ 64
 DB 31 RSNALNDYQSRKGLVLMNNG-GVKQEIQSIQVCAKQCS-----LLDCRSFYNNKKSQT 85
 QY 65 CLWFPNSMSSGVKKEFEHFDLYENKDYIRNCLIGKGRSYKGTVSITKSGIKCOPWSSM 124
 DB 86 CRLLPWTQNSANVLLQNVQYDLQKDYIRDCVAGNGTYRTVTSKSGRTQCMRLK 145
 QY 125 IPHEHFLPSSYRGKDLQENYCRNPRGEGGPMCFSTSNPEVRYEVCIDIPQCSVEVCMTCN 184
 DB 146 FPDHDFSLHW--PELENYCRNPDSDPEGWCYTTDNIRHOYCGIKKCEDAVCLTCN 203
 QY 185 GSYRGMDHTESGKTCQRMWDHQTPIRHKFLPERYPDKGFDNNYCRNPDGQPRPWCYTLD 244
 DB 204 GEDYRGFVDTESGTCQRMWDLQHPHKKHPYHPDKYPEKGLDNNYCRNPDSSSEPMCYTTD 263
 QY 245 PHTRWYCAIKTCADNTMDTVPLETTCTCQGGEGYRGVTWNTIWIWGPCQWDSQYP 304
 DB 264 PNVKEFCIRVCKKPR-----TSTCFKRGEGYRGKANTTSGIPQQRWDSQTPQ 320
 QY 305 EHDMTPEKCKDLRENYCRNPDGSEPMCFSTDPNIRVGYCSQIPNCDMS-HGQDCYR 363
 DB 321 SHRFLPEKYPCKDLQENYCRNPDGSEAPWCFTTRGMRVAFCFHTRCDDDELDAECYH 380
 QY 364 GNGKNTWGNLSQTRSGLTCSMDKNMEDLHRHI-----FWEPDASKLNYCRNPDODDAH 423
 DB 381 NGELYSGRYSKTRKGIKCRWEKRNDELSDL-QPYLVPLEENYCRNPDSDSHGWCYT 439
 QY 424 GNPILPWDCYPSRCEGDTTPTI 446
 DB 440 MDPNTDFDYCAIKPCSGEKVLT 462

RESULT 9
 P70006 PRELIMINARY; PRT; 717 AA.
 AC P70006;
 DT 01-FEB-1997 (TremBLrel. 02, Created)
 DT 01-FEB-1997 (TremBLrel. 02, Last sequence update)
 DT 01-MAR-2002 (TremBLrel. 20, Last annotation update)
 DE Hepatocyte growth factor-like protein precursor.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
 OC Xenopodinae; Xenopus.
 OX NCBI_TaxID=8355;
 RN [1]

RP SEQUENCE FROM N.A.
 RC TISSUE-KIDNEY;
 RX MEDLINE=96404125; PubMed=8808403;
 RA Aberger F., Schmidt G., Richter K.;
 RT "The xenopus homologue of hepatocyte growth factor-like protein is
 RT specifically expressed in the presumptive neural plate during
 RT gastrulation.";
 RL Mech. Dev. 54:23-37(1996).
 DR EMBL; Y08734; CAA69989.1; -;
 DR HSSP; P00747; ICPA.
 DR WEROPS; S01.977; -;
 DR InterPro; IPR000001; Kringle.
 DR InterPro; IPR003014; PAN.
 DR InterPro; IPR003609; Pan_app.
 DR InterPro; IPR001254; Ser_protease_Try.
 DR Pfam; PF00051; kringle; 4.
 DR Pfam; PF00024; PAN; 1.
 DR Pfam; PF00089; trypsin; 1.
 DR PRINTS; PR00018; KRINGLE.
 DR ProDom; PD000395; kringle; 4.
 DR SMART; SM00130; KR; 4.
 DR SMART; SM00473; PAN_AP; 1.
 DR SMART; SM00020; TRYP_SPC; 1.
 DR PROSITE; PS00021; KRINGLE_1; 3.
 DR PROSITE; PS00070; KRINGLE_2; 4.
 DR PROSITE; PS00240; TRYP_SIN_DOM; 1.
 KW Hydrolase; Serine protease; Signal.
 FT SIGNAL 1 28
 FT CHAIN 29 717
 FT SEQUENCE 717 AA; 82017 MW; 6F877A432C8CDD54 CRC64;

Query Match 45.2%; Score 1176; DB 13; Length 717;
 Best Local Similarity 45.4%; Pred. No. 7.5e-101;
 Matches 201; Conservative 78; Mismatches 152; Indels 12; Gaps 6;

QY 5 RNTTHERKSAKTLIKIDPALKIKTKVNTADQCANRCTRNGLPFTCKAFVDFKARKQ 64
 Db 32 RSLANDYQSRKGLVHNEG-GVKQEVQAEIQICAKQCSQD---LLDCRSFDYNNKSQS 86
 QY 65 CLWPFNMSGSGVKKFEGHFDLYENKDYIRNCLIGKGRSYKGTSTKSGIKQCPWSSM 124
 Db 87 CRLLPWTQNSNVLLRNQVYDLYKKDYIRDCVVGNGNTYRGTVSKTKNGRTCCQHWRLK 146
 QY 125 IPHSHSLPSSYRGKDLQENYCRNPRGEGGPGWCTSNPEVRYEVCDDIPQCSVEECMTCN 184
 Db 147 FPHBKESPTHW--PELEENYCRNPDSDPEGLWCYTDDKNIHQYCGIKKCEDAVCLTCN 204
 QY 185 GESYRGMLDHTESGKICQWHDQPHRHKFLPERYPDKGFDNNYCRNPDQCPRPWCYTLD 244
 Db 205 GEDYRGSDRTESGKQWDLQAPHTHPYKPEKYPDKSLDDNYCRNPDSSERPWCYTLD 264
 QY 245 PHTRWYCAIKTCADNTMNDVDVPLETTECIGOGEGYRGTVNTIANGIPCQRWDSQYPH 304
 Db 265 PNVEREFCHITCKQRISNIEI---TSTCFKGEGRGKANTTTSGIPCQRWDIOAPH 321
 QY 305 EHDMTPEFNKCDLRENYCRNPDGSESPWCTTDPNTRVGYCSQIPNC-DMSGQDCYRG 363
 Db 322 VHRFLPEYPCCKGLDENYCRNPDGSEAPWCTTTLKNRMAYCFQIKRCTDDVVEPECYHG 381
 QY 364 NGKNTYGNLSQTRSLGTCMSMDKNMEDLHRHIFWEPDASKLNENYCRNPDGDAHGPWCYT 423
 Db 382 NGELYRGVSKTRKGMCRWRWEEKNDLELSLA-OPYLVPLEENYCRNPDSDSHGPWCYT 440
 QY 424 GNPLIPWPCYPSRCEGDTPTTI 446
 Db 441 MDPNTPFDYCALKPCAGDKVLT 463

RESULT 10
 Q90ZNG PRELIMINARY; PRT; 709 AA.
 ID Q90ZNG
 AC Q90ZNG
 DT 01-DEC-2001 (Tremblrel. 19, Created)

DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
 DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
 DE Hepatocyte growth factor-like 1.
 GN HGFLI.
 OS Brachydanio rerio (Zebrafish) (Zebra danio).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
 OC Cyprinidae; Danio.
 OX NCBI_TaxID=7955;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Bassett D.I., Wilson S.W.;
 RT "Early expression of zebrafish Hepatocyte Growth Factor-Like 1
 RT suggests a conserved role in vertebrate neural induction.";
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF370035; AAK54207.1; -;
 DR InterPro; IPR000001; Kringle.
 DR InterPro; IPR003014; PAN.
 DR InterPro; IPR001254; Ser_protease_Try.
 DR Pfam; PF00051; kringle; 4.
 DR Pfam; PF00024; PAN; 1.
 DR Pfam; PF00089; trypsin; 1.
 DR ProDom; PD000395; kringle; 4.
 DR PROSITE; PS00021; KRINGLE_1; UNKNOWN_4.
 DR PROSITE; PS00070; KRINGLE_2; 4.
 DR PROSITE; PS00240; TRYP_SIN_DOM; 1.
 KW Hydrolase; Serine protease.
 SQ SEQUENCE 709 AA; 81271 MW; 9907236C5DB73A20 CRC64;

Query Match 45.0%; Score 1172; DB 13; Length 709;
 Best Local Similarity 46.0%; Pred. No. 1.8e-100;
 Matches 216; Conservative 60; Mismatches 134; Indels 60; Gaps 12;

QY 5 RNTTHERKSAKTLIKIDPALKIKTKVNTADQCANRCTRNGLPFTCKAF 56
 Db 21 RSLANDYQSRKGLVHNEG-GVKQEVQAEIQICAKQCSQD---LLDCRSFDYNNKSQS 86
 QY 57 VFD-KARKQCLWFPF--NSMSGVKKFEGHFDLYENKDYIRNCLIGKGRSYKGTSTK 113
 Db 70 NYEFPRLSVCKHLPWVGDNADVRNV--NCGLYEMKVYVRKCIYVGKEDYRGKVSTTT 127
 QY 114 SGIKQCPWSSMIPHEHSLPSSYRGKDLQENYCRNPRGEGGPGWCTSNPEVRYEVCDDIP 173
 Db 128 SGRTQQQWMSKPHDHRWTPSATNGLEL--NCRNPDGDRIGPWCYTDPERYESCNIP 185
 QY 174 QCSVEECMTGNGESYRGMLDHTESGKICQWHDQPHRHKFLPERYPDKGFDNNYCRNPD 233
 Db 186 QCKDEVCITCNGEDYRGQVDHTISGKQWDLQAPHTHPYKPEKYPDKSLDDNYCRNPD 245
 QY 234 GQPRPWCYTLDPHTRWYCAIKTCADNTMNDVDVPLETTECIGOGEGYRGTVN 287
 Db 246 ASPVPWCYTDPDPTMERESCDISK-----PEFPRLRSLRSYTTNCFAGGEDYRGKN 298
 QY 288 TIWNGIPCQRWDSQYPHEDMTPEFNKCDLRENYCRNPDGSESPWCTTDPNTRVGYCS 347
 Db 299 EYTSIGIPCQRWDSQYPHEDMTPEFNKCDLRENYCRNPDGSESPWCTTDPNTRVGYCS 358
 QY 348 QIPNC-DMSGQDCYRGKNTYGNLSQTRSLGTCMSMDKNMEDLHRHIFWEPDASKLN- 405
 Db 359 QIKRCADDIEADCYNEIGKNYRGVVRKTRKGLCQKWSIN-----TPHKTKNP 408
 QY 406 -----ENYCRNPDGDAHGPWCYTGNPLIPWPCYPSRCEGDTPTTI 446
 Db 409 KTHPEANLTDNYCRNPDGDHGHGPWCYTSPDKTEFYCALKQACAGEKVP 458

RESULT 11
 Q13208 PRELIMINARY; PRT; 567 AA.
 ID Q13208
 AC Q13208
 DT 01-NOV-1996 (Tremblrel. 01, Created)
 DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
 DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)

DE Hepatocyte growth factor-like protein homolog.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20191171; PubMed=10728827;
 RA Degen S.J.F., McDowell S.A., Waltz S.E., Gould F., Stuart L.A.,
 RA Carritt B.;
 RT "Structure of the human D1F1S1A locus: a chromosome 1 locus with 97%
 RT identity to the chromosome 3 gene coding for hepatocyte growth factor-
 RT like protein.";
 RL DNA Seq. 8:409-413(1998).
 DR EMBL; U28054; AAC63092.1; -.
 DR HSSP; P00747; 2PK4.
 DR MEROPS; S01.977; -.
 DR InterPro; IPR000001; Kringle.
 DR InterPro; IPR003014; PAN.
 DR InterPro; IPR003609; Pan_app.
 DR Pfam; PF00051; kringle; 4.
 DR Pfam; PF00024; PAN; 1.
 DR Pfam; PF00089; trypsin; 1.
 DR PRINTS; PR00018; KRINGLE.
 DR PRODOM; PD000395; Kringle; 4.
 DR SMART; SM00130; KR; 4.
 DR SMART; SM00473; PAN_AP; 1.
 DR PROSITE; PS00021; KRINGLE_1; 2.
 DR PROSITE; PS00070; KRINGLE_2; 4.
 DR PROSITE; PS02040; TRYPSIN_DOM; 1.
 DR Hydrolase; Serine protease.
 KW Hydrolyase; Serine protease.
 SQ SEQUENCE 567 AA; 64117 MW; 3FC38B07F1645810 CRC64;
 Query Match 44.78; Score 1163.5; DB 4; Length 567;
 Best Local Similarity 47.18; Pred. No. 8.2e-100;
 Matches 196; Conservative 62; Mismatches 149; Indels 9; Gaps 4;
 Qy 33 VNTADQCANRTRKGLPFTCKAFVDFKARKQCLWFPNSMSGVKKFEGHEFDLYENKD 92
 Db 25 VADAEACAGRC---GLLMDCAWFHYNVSSHGCCQLLPWQHSPHSLRHSRGCDLFOKDD 80
 Qy 93 YIRNCITGKSGYKGTVITKSGIKQCPWSSMIPHEHSFLPSSYRGKDLQENYCRNPRGE 152
 Db 81 YIRTCIMNGVGYRDTMTATVGLSCQAWSHKFPNDHQYMTPLRNG--LEENFCRNPDGD 138
 Qy 153 EGGPWCFSTSNPEVRYEYCDIPQCEVECMTCNGESYRGKLDHTEGSKICQWDHQTPIRH 212
 Db 139 PGGPWCHTTPDAVRFSQGIKSLVAACVWCNGEYRGAVDRTESGREGCQWDLQHPHQH 198
 Qy 213 KFLPERYPDKGDFDNYCRNPDGQPRPWCYTLDPHTRWEYCAIKTCADNTMDTDVPLETT 272
 Db 199 PFGKFLDQGLDDNYCRSPDGSQRPWCYTTPQIEREFCDLPRGSEAQPRQEA--TSV 256
 Qy 273 ECIQGGEGYRGTVNTIWTNGIPCORWDSQYPEHDMTPENFKCKDLRNYCRNPDGSGSP 332
 Db 257 SCFRGKGEGYRGTVNTIWTAGVPCQWDAQIPQHRTPEKTYACKDLRNFENCRNPDGSEAP 316
 Qy 333 WCFTDPNIRVGYCSQIPNC--DMSGHGDQCYRGNGKNYGNLSQTSGLTCSMDKKNMEDL 391
 Db 317 WCFTLPRTGRFVCFQIRRCFDDVRPDQCYHGAGEYQYKTSKTRKGVQOCQWSAETPHK 376
 Qy 392 HRIHFWEPDASKLNYCRNPDGDAHPWCYTGNPLIPWDYCPISRCBGDTPTTIV 447
 Db 377 PQFTFTSEPHQAQLEENFCQDGDGSHGWCYTMDPRTFFDYCALRCADDQPPSIL 432
 RESULT 12
 ID P70521 PRELIMINARY; PRT; 716 AA.
 AC P70521;
 DT 01-FEB-1997 (TrEMBLrel. 02, Created)
 DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)

DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE Macrophage stimulating protein precursor.
 GN MSP.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97011126; PubMed=8858136;
 RA Onshiro K., Iwama A., Matsuno K., Ezaki T., Sakamoto O., Hamaguchi I.,
 RA Takasu N., Suda T.;
 RT "Molecular cloning of Rat Macrophage-stimulating protein and its
 RT involvement in the Male Reproductive System.";
 RL Biochem. Biophys. Res. Commun. 227:273-280(1996).
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
 CC TRYPSIN FAMILY.
 DR EMBL; X95096; CAA64473.1; -.
 DR HSSP; P00747; 1KRN.
 DR MEROPS; S01.975; -.
 DR InterPro; IPR001314; Chymotrypsin.
 DR InterPro; IPR000001; Kringle.
 DR InterPro; IPR003014; PAN.
 DR InterPro; IPR003609; Pan_app.
 DR Pfam; PF00051; kringle; 4.
 DR Pfam; PF00024; PAN; 1.
 DR Pfam; PF00089; trypsin; 1.
 DR PRINTS; PR00722; CHYMOTRYPSIN.
 DR PRINTS; PR00018; KRINGLE.
 DR PRODOM; PD000395; Kringle; 4.
 DR SMART; SM00130; KR; 4.
 DR SMART; SM00473; PAN_AP; 1.
 DR SMART; SM00020; Tryp_Spc; 1.
 DR PROSITE; PS00021; KRINGLE_1; 4.
 DR PROSITE; PS00070; KRINGLE_2; 4.
 DR PROSITE; PS02040; TRYPSIN_DOM; 1.
 DR Hydrolase; Serine protease; Signal.
 KW SIGNAL; 31 POTENTIAL.
 SQ SEQUENCE 716 AA; 80733 MW; 06B7DF3EF56D921F CRC64;
 Query Match 43.9%; Score 1144; DB 11; Length 716;
 Best Local Similarity 46.18; Pred. No. 7.2e-98;
 Matches 196; Conservative 70; Mismatches 141; Indels 18; Gaps 6;
 Qy 33 VNTADQCANRTRKGLPFTCKAFVDFKARKQCLWFPNSMSGVKKFEGHEFDLYENKD 92
 Db 50 VADAEACAGRC---GPLLDCRAFHYNMSSHGCCQLLPWQHSLRAQLHSSLCDLFOKDD 105
 Qy 93 YIRNCITGKSGYKGTVITKSGIKQCPWSSMIPHEHSFLPSSYRGKDLQENYCRNPRGE 152
 Db 106 YVTCIMDNGASYRGTVARTADGLPCQAWSRFRFNDHKYTPPKNG--LEENFCRNPDGD 163
 Qy 153 EGGPWCFSTSNPEVRYEYCDIPQCEVECMTCNGESYRGKLDHTEGSKICQWDHQTPIRH 212
 Db 164 PRGPWCYTNNRVSFQSCGIKSCREAVCWCNGEDYRGEDVDTESGREGCQWDLQHPHSH 223
 Qy 213 KFLPERYPDKGDFDNYCRNPDGQPRPWCYTLDPHTRWEYCAIKTCADN-----TMDTD 265
 Db 224 PFHPEKFPDKALKONCRNPDASERPWCYTTPDNVEREFCDLPSGPNLPPTTKGSKSQ 283
 Qy 266 DVPLETTCIOGEGYRGTVNTIWTNGIPCORWDSQYPEHDMTPENFKCKDLRNYCRN 325
 Db 284 RNKYKASNCFRGKGEDYRGTVNTTSAGVPCQWDAQNPHQHFVPEKYACKDLRNFENCRN 343
 Qy 326 PDGSESPCFCTDPNIRVGYCSQIPNC--DMSGHGDQCYRGNGKNYGNLSQTSGLTCSMW 384
 Db 344 PDGSEAPWCFTSRPGLRVAFCYQIPRCTEEVVPVPECVYHGSQEQYRGVSKTRKGVQOCQHW 403
 Qy 385 DKNMEDLHRIHFWEPDA--SKLNYCRNPDGDAHPWCYTGNPLIPWDYCPISRCBGDT 442
 Db 404 --SSETPHKPQFTPTSAPHAGLEANFCRNPDGDSHGWCYTLDPETLFDYCALKRCDDQ 461

Db 462 PPSIL 466

RESULT 14

Q9ROW3 PRELIMINARY: PRT: 812 AA.

ID Q9ROW3

AC 01-MAY-2000 (TReMBLrel. 13, Created)

DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)

DT 01-MAY-2002 (TReMBLrel. 20, Last annotation update)

DT 01-MAY-2002 (TReMBLrel. 20, Last annotation update)

DE Plasmidogen protein precursor (EC 3.4.21.7).

GN PLASMINOGEN

OS Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.

OX NCBI_TaxID=10116;

OX [1]

SEQUENCE FROM N.A.

RC TISSUE=LIVER;

RC TISSUE=LIVER;

RA Banget K., Johnsen A.H., Thorsen S.;

RT "Rat plasminogen: cDNA and gene structure."

RT Submitted (MAY-1999) to the EMBL/GenBank/DBAJ databases.

RN [2]

SEQUENCE FROM N.A.

RP TISSUE=LIVER;

RP MEDLINE=91250378; PubMed=1645711;

RC MEDLINE=91250378; PubMed=1645711;

RC Kanalas J.J., Makker S.P.;

RA "Identification of the rat Heymann nephritis autoantigen (GP330) as a

RT receptor site for plasminogen."

RT J. Biol. Chem. 266:10825-10829 (1991).

CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE

CC TRYPSIN FAMILY.

CC EMBL; AJ242649; CAB46014.1; -.

DR HSSP; P00747; 1PMK.

DR MEROPS; S01.233; -.

DR InterPro; IPR001314; Chymotrypsin.

DR InterPro; IPR000001; Kringle.

DR InterPro; IPR003014; PAN.

DR InterPro; IPR003609; Pan_app.

DR InterPro; IPR001254; Ser_protease_Try.

DR InterPro; IPR001400; Somatotropin.

DR Pfam; PF00051; kringle; 5.

DR Pfam; PF00024; PAN; 1.

DR Pfam; PF00089; trypsin; 1.

DR PRINTS; PR00722; CHYMOTRYPSIN.

DR PRINTS; PR00018; KRINGLE.

DR ProDom; PD000395; Kringle; 5.

DR SMART; SM00130; KR; 4.

DR SMART; SM00473; PAN_AP; 1.

DR SMART; SM00020; TYP_SPC; 1.

DR PROSITE; PS00021; KRINGLE_1; 5.

DR PROSITE; PS00070; KRINGLE_2; 5.

DR PROSITE; PS00338; SOMATOTROPIN_2; UNKNOWN_1.

DR PROSITE; PS0240; TRYPSIN_DOM; 1.

DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.

DR PROSITE; PS00135; TRYPSIN_SER; 1.

KW Hydrolase; Serine protease; Signal.

FT SIGNAL 1 19

FT CHAIN 20 812 PLASMINOGEN.

FT SEQUENCE 812 AA; 90535 MW; 8C703C51410EBC9E CRC64;

Query Match 41.7%; Score 1085.5; DB 11; Length 812;

Best Local Similarity 45.0%; Pred. No. 2.3e-92;

Matches 193; Conservative 65; Mismatches 140; Indels 31; Gaps

Qy 30 TKKVTADQCANRTRNGK-LPFTCKAFVDFDKARQKCLWFFNSMSSGVKKEFGHEEDLY 88

Db 37 TKKLAAGSIADCLAKCEGDETFDCRSFQVHSKQCCQVMAENSKTSIIIRM--RDVILF 94

Qy 89 ENKQYIRNCIIAGKRSYKGTVSITKSGIKTOPWSESMIPHEHSFLPSSYRGKDLQENYCRN 148

Db 95 EKRVLSECKTGIGKYGRTGSKTKGTGTCQKWSDTSPHPVKYSPSTHPSEGLEENYCRN 154

QY 149 PRGEGGWCFTSNPEVRYEVCIDPOCSEVECMTCNGESYRGLMDHTESGKTCQWRDHT 208
 Db 155 PNDQGGWCYTTDPQRYEYCNIECEB-ECMYSGEYEGKISKTMISGLDCQSWDSOS 213
 QY 209 PHRHKFLPERYDKGFDNDYCNPNQDGPWCYTLDPHTRWEYCAIKTCADNTMNDTVP 268
 Db 214 PHAGYIIFAKFSKMLKMYCNPNQDGPWCYTLDPHTRWEYCAIKTCADNTMNDTVP 269
 QY 269 LETTEICQOGGYRGTVNTIWNIGIPQORWDSOYRPHEDMTENFKCKDLNRYCNPNQD 328
 Db 270 GFTYOCCKGRGENYRGTVSVTASGKTCQWRSEOTPHRHRTNPEFCKNLEENYCNPNQD 329
 QY 329 SESPCFTTDPNIRVYCSQIPNC-----DMSHG-----ODCYRGNGKMYMGNL 372
 Db 330 ETAPWCYTTDSLRWEYC-EIPSCSSVSPDQSSVLPPEQTPVQVQECYQNGKSYRGTS 388
 QY 373 SOTRSGLTCSMDKWNEDLHRHI---FWEPDASKLENYCNPNQDADGAPWCYTCNPLIP 429
 Db 389 SYTNTGKCKQSVSWTPHSHSKTANF--PDAG-LEMNYCNPNQDQDGPWCFTTDPDSVR 445
 QY 430 WDCYPISRC 438
 Db 446 WEYCNLKR 454

RESULT 15

OL8783
 ID OL8783 PRELIMINARY; PRT; 806 AA.
 AC OL8783;
 DT 01-JAN-1998 (Tremblrel. 05, Created)
 DT 01-JAN-1998 (Tremblrel. 05, Last sequence update)
 DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
 DE Plasminogen.
 OS Macropus eugenii (Tamar wallaby).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Metatheria; Diprotodontia; Macropodidae; Macropus.
 OX NCBI_TaxID=9315;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=LIVER;
 RX MEDLINE=98004511; PubMed=9342350;
 RA Lawn R.M., Schwartz K., Pathy L.;
 RT "Convergent evolution of apolipoprotein(a) in primates and hedgehog."
 RL Proc. Natl. Acad. Sci. U.S.A. 94:11992-11997(1997).
 CG -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
 CC TRYPSIN FAMILY.
 DR EMBL; AF012297; AAB65760.1; -;
 DR HSP; P00747; 5HPG.
 DR MEROPS; S01.233;
 DR InterPro; IPR001314; Chymotrypsin.
 DR InterPro; IPR000001; Kringle.
 DR InterPro; IPR003014; PAN.
 DR InterPro; IPR003609; Pan_app.
 DR InterPro; IPR001254; Ser_protease_Try.
 DR Pfam; PF00051; kringle; 5.
 DR Pfam; PF00024; PAN; 1.
 DR Pfam; PF00089; trypsin; 1.
 DR PRINTS; PR00722; CHYMOTRYPSIN.
 DR PRINTS; PR00018; KRINGLE.
 DR ProDom; PD000395; Kringle; 5.
 DR SMART; SM00130; KR; 4.
 DR SMART; SM00473; PAN_AP; 1.
 DR SMART; SM00020; Tryp_SPC; 1.
 DR PROSITE; PS00021; KRINGLE.1; 5.
 DR PROSITE; PS00070; KRINGLE.2; 5.
 DR PROSITE; PS0240; TRYPSIN_DOM; 1.
 DR PROSITE; PS00134; TRYPSIN_HIS; 1.
 DR PROSITE; PS00135; TRYPSIN_SER; 1.
 KW Hydrolase; Serine protease.
 SQ SEQUENCE 806 AA; 90981 MW; 95FAA86DC20064D5 CRC64;

Query Match 41.5%; Score 1080.5; DB 6; Length 806;
 Best Local Similarity 42.8%; Pred. No. 6.8e-92;

Matches 191; Conservative 79; Mismatches 143; Indels 33; Gaps 11;
 QY 20 IKIDPALKTKTKK-----VNTADQCANRCTRNKGLPETCKAFVDFDKARKOCLAFPPFNMS 75
 Db 26 IKTEGASISNSOKKOFVASSTEECALC--EKETEVCRSFEHYNKQKCVIMSENKTS 83
 QY 76 GYKKEFGHEFDLYENKDYIRNCIIIGKGRSYKGTVSITKSGIKCQPMSSMIPHEHSLPSS 135
 Db 84 SVERRK--RDVLFKEKRIYLSLDCSKSGNGRNYRGLTSLKTSKITCQKWSLSPHPVNPVPSK 141
 QY 136 YGKDLQENYCNPRGEGGWCFTSNPEVRYEVCIDPOCSEVECMTCNGESYRGLMDHT 195
 Db 142 YPDAGLEKNYCNPNQDQDGPWCYTTNPDRIYECVPEEC-EDCEMHCSENGYRGITSKT 200
 QY 196 ESGKICQWRDHTPHRHKFLPERYDPKGFDDNVCNPNQDGPWCYTLDPHTRWEYCAIK 255
 Db 201 ESGIECPWDSOEPHSHEYIPSKFSPKDLKENYCNPNQDGPWCFTSNPEKWEFCNP 260
 QY 256 TCADNTMNDTVDPLETTEICQOGGYRGTVNTIWNIGIPQORWDSOYRPHEDMTENFKC 315
 Db 261 RCS----SPPPPPGPMQLCKLGRGENYRGKIAVTKSGHTCQRWKNKQTPHKHRTPENFPC 316
 QY 316 KDLRENYCNPNQDSESPWCFTTDPNIRVYCSQIPNCDSHG-----ODCYRGN 364
 Db 317 RGLDENYCNPNQDGELEPWCYTTNPDVQYCA-IPSCGTSSPHTDRVEQSPVIOECYEKG 375
 QY 365 GKNYMGNLISOTRSGLTCSMDKWNEDLHRHIFWEPD---ASKLNEYCNPNQDADGAPWC 421
 Db 376 GENYRGTVTSTTISGKKCQAWSSMTPHQKK---TPDNFFNADLIRNYCNPNQD-KSPWC 431
 QY 422 YTGNLIPWDYCPISRCGDTTPTIV 447
 Db 432 YTMDDTVRWEFCNLEKCSG-TGSTVL 456

Search completed: April 25, 2003, 15:51:43
 Job time : 40 secs

THIS PAGE BLANK (USPTO)